Automatic Extraction of Cardiac Time Intervals From Synchronous ECG and PCG

Mariana Lourenço
Master in Data Science
Department of Computer Science
Faculty of Sciences of the University of Porto
2025

Automatic Extraction of Cardiac Time Intervals From Synchronous ECG and PCG

Mariana Lourenço
Dissertation carried out as part of the Master in Data Science
Department of Computer Science
2025
Supervisor
Francesco Renna, Assistant Professor, Faculty of Sciences of the University of Porto

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

INESCTEC
TECHNOLOGY & SCIENCE
ASSOCIATE LABORATORY
PORTUGAL

Acknowledgements

Gostaria de expressar a minha profunda gratidão ao meu orientador, Professor Francesco Renna, pelo apoio, orientação e disponibilidade ao longo deste percurso. As suas diretrizes no meu trabalho foram essenciais para a concretização deste projeto.

Um agradecimento especial ao colega de doutoramento Daniel Proãno, pela ajuda, pelas dúvidas tiradas e pela companhia durante esta jornada.

Aos meus pais e ao meu irmão, o meu obrigado pelo apoio incondicional, paciência e incentivo constante.

Às minhas amigas, Inês Tavares e Carolina Amaral, agradeço pelo carinho, motivação e longas noites de estudo juntas.

Resumo

A insuficiência cardíaca (IC) afeta mais de 64 milhões de pessoas em todo o mundo e representa uma carga significativa para os sistemas de saúde. A avaliação precisa da função cardíaca é essencial para o acompanhamento e tratamento da IC, sendo a ecocardiografia o padrão-ouro clínico para a medição de intervalos de tempo cardíacos (CTIs), como o Período Pré-Ejetivo (PEP) e a Fração de Ejeção do Ventrículo Esquerdo (LVEF). No entanto, o elevado custo e a necessidade de pessoal especializado limitam o acesso a esta técnica. Esta dissertação investiga métodos alternativos e não invasivos para estimar os CTIs a partir de registos de eletrocardiograma (ECG) e fonocardiograma (PCG). O ECG capta a atividade elétrica do coração, enquanto o PCG regista os sons cardíacos associados aos movimentos mecânicos das válvulas. O trabalho inclui o desenvolvimento de um pipeline Beat-by-Beat para segmentação automática dos pontos fiduciais no ECG e PCG, bem como a proposta de uma abordagem robusta baseada no Batimento Cardíaco Mediano (Median Heartbeat) para melhorar a estimativa dos CTIs em sinais de baixa qualidade. Ambos os métodos são comparados entre si e validados com base em anotações de referência. Adicionalmente, é analisada a relação entre os CTIs estimados e os grupos de pacientes estratificados pela LVEF. Os resultados obtidos reforçam a viabilidade do uso combinado de ECG e PCG para monitorização acessível e de baixo custo da função cardíaca, especialmente em contextos onde técnicas de imagem convencionais não estão disponíveis.

Palavras-chave: Eletrocardiograma, Fonocardiograma, Intervalos de tempo cardíacos, Período pré-ejetivo, Tempo de ejeção ventricular esquerda, Fração de ejeção do ventrículo esquerdo, Processamento de sinais biomédicos, U-Net

Abstract

Heart Failure (HF) affects over 64 million people and places a heavy burden on healthcare systems. Accurate evaluation of cardiac function is crucial for HF management, and although echocardiography is the clinical gold standard for measuring cardiac time intervals (CTIs) like the Pre-Ejection Period (PEP) and Left Ventricular Ejection Fraction (LVEF), its high cost and need for specialized personnel limit its accessibility. This thesis investigates alternative, non-invasive methods for estimating CTIs using electrocardiogram (ECG) and phonocardiogram (PCG) recordings. ECG captures electrical cardiac activity, while PCG records heart sounds associated with mechanical valve movements. The work involves developing a Beat-by-Beat pipeline for segmenting ECG and PCG fiducial points, and a robust Median Heartbeat Approach to improve CTI estimation in low-quality recordings, comparing both methods, and validating their estimates against ground truth annotations. Additionally, it examines how the estimated CTIs relate to patient groups stratified by LVEF. The findings support the feasibility of using ECG and PCG for accessible, cost-effective monitoring of cardiac function, particularly in contexts where traditional imaging is not available.

Keywords: Electrocardiogram, Phonocardiogram, Cardiac time intervals, Pre-ejection period, Left ventricular ejection time, Left ventricular ejection fraction, Biomedical signal processing, U-Net

Table of Contents

List of Figures

List of Abbreviations

CTIs Cardiac Time Intervals. iv, v, 1– 3, 5, 6, 10– 15, 17, 23, 25, 26, 41, 42, 50, 55, 57

CVDs Cardiovascular diseases. 1

ECG Electrocardiogram. vi, 1– 3, 5– 7, 10– 13, 24, 26– 31, 34, 37– 40

ECHO Echocardiography. 1, 4, 5

EDV End-Diastolic Volume. 4

EMAT Electromechanical Activation Time. 10– 12

HF Heart Failure. vi, 1, 3– 5, 12

HS Heart Sound. 1

IVCP Isovolumic Contraction Period. 10

LVEF Left Ventricular Ejection Fraction. 1– 5, 12, 24, 25, 58

LVET Left Ventricular Ejection Time. vi, 1, 10– 12, 17– 22, 42

LVST Left Ventricular Systolic Time. 10– 12

ML Machine Learning. 1

PCG Phonocardiogram. vi, 1– 3, 5, 6, 10– 13, 24, 26– 30, 38– 40

PEP Pre-Ejection Period. vi, 1, 10– 12, 17– 22, 42

QS2 Total Electromechanical Systole. 10, 42

RR Time Interval Between Successive R Waves. 11

S1 First Heart Sound. 1, 10, 25

S2 Second Heart Sound. 1, 10, 25

SCICA Semi-blind Canonical Independent Component Analysis. 21

1. Introduction

1.1. Overview

Cardiovascular diseases (CVDs) are the leading cause of death globally, accounting for an estimated 17.9 million deaths in 2019, which represents 32% of all deaths worldwide [1]. Among CVDs, Heart Failure (HF) remains a significant public health challenge, affecting over 64 million people worldwide and imposing a substantial burden on healthcare resources [2]. Accurate assessment of cardiac function is critical for the diagnosis and management of HF. Echocardiography (ECHO) is considered the clinical gold standard for measuring Cardiac Time Intervals (CTIs) such as the Pre-Ejection Period (PEP) and Left Ventricular Ejection Fraction (LVEF), which are vital indicators of systolic performance [3]. However, ECHO requires expensive equipment and highly trained personnel, limiting its accessibility in many settings [4].

Alternatively, non-invasive cardiac monitoring modalities such as the Electrocardiogram (ECG) and Phonocardiogram (PCG) offer cost-effective and widely available means to estimate CTIs. ECG wide use in cardiac diagnostics can be attributed to its simplicity, non-invasiveness, and low cost [5]. PCG records the mechanical vibrations of Heart Sound (HS) and provides information on valve opening and closure, enabling estimation of PEP and Left Ventricular Ejection Time (LVET) from the timing of the First Heart Sound (S1) and Second Heart Sound (S2) HS [6]. Recent advances in signal processing and Machine Learning (ML) have further enhanced the feasibility of extracting these systolic time intervals from combined ECG and PCG signals, potentially enabling continuous monitoring of HF patients [7]. Despite these advances, it is essential to validate the accuracy and clinical utility of ECG- PCG derived measurements.

1.2. Objectives

The primary objective of this thesis is to develop and validate methods for estimating CTIs using non-invasive ECG and PCG recordings. In line with the **SMART** criteria (Specific, Measurable, Achievable, Relevant, Time-bound), the specific aims are:

- **Specific & Measurable:** Develop a beat-by-beat pipeline to segment ECG and PCG fiducial points, enabling automatic CTIs extraction with accuracy evaluated against annotated datasets.

- **Specific & Achievable:** Propose a signal pre-processing step (the Median Heart-beat approach) to extract CTIs from low-quality recordings.
- **Measurable & Relevant:** Compare the performance of the Beat-by-Beat extraction method and the Median Hearbeat approach using quantitative error metrics (e.g., mean absolute error, Bland–Altman limits of agreement).
- **Measurable & Relevant:** Validate ECG- PCG derived CTIs estimates against manual annotations taken from these recording.
- **Relevant & Time-bound:** Investigate the correlation between ECG- PCG derived CTIs and patient groups stratified by Left Ventricular Ejection Fraction, completing the analysis within the thesis project timeline.

2. Background

This chapter establishes the context for the problem under investigation. It begins by introducing Heart Failure and its critical association with Left Ventricular Ejection Fraction, which serves as a fundamental indicator of cardiac function.

The discussion then turns to Cardiac Time Intervals measurements that capture the timing of mechanical and electrical cardiac events. These intervals have been shown to offer discriminative power in differentiating between patient groups with varying levels of LVEF, thereby providing a means to assess cardiac performance.

Finally, we highlight how it is possible to identify the physiological events in synchronized ECG and PCG signals. The precise identification of these fiducial points enables accurate estimation of key intervals, offering valuable insights into the electromechanical function of the heart.

2.1. Heart Failure and Left Ventricular Ejection Fraction

When a patient is suspected to have HF, diagnostic tests are run to confirm or exclude the HF diagnosis. The diagram in figure 2.1 presents the diagnostic algorithm proposed by the European Society of Cardiology (ESC) in the "2023 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure" [8].

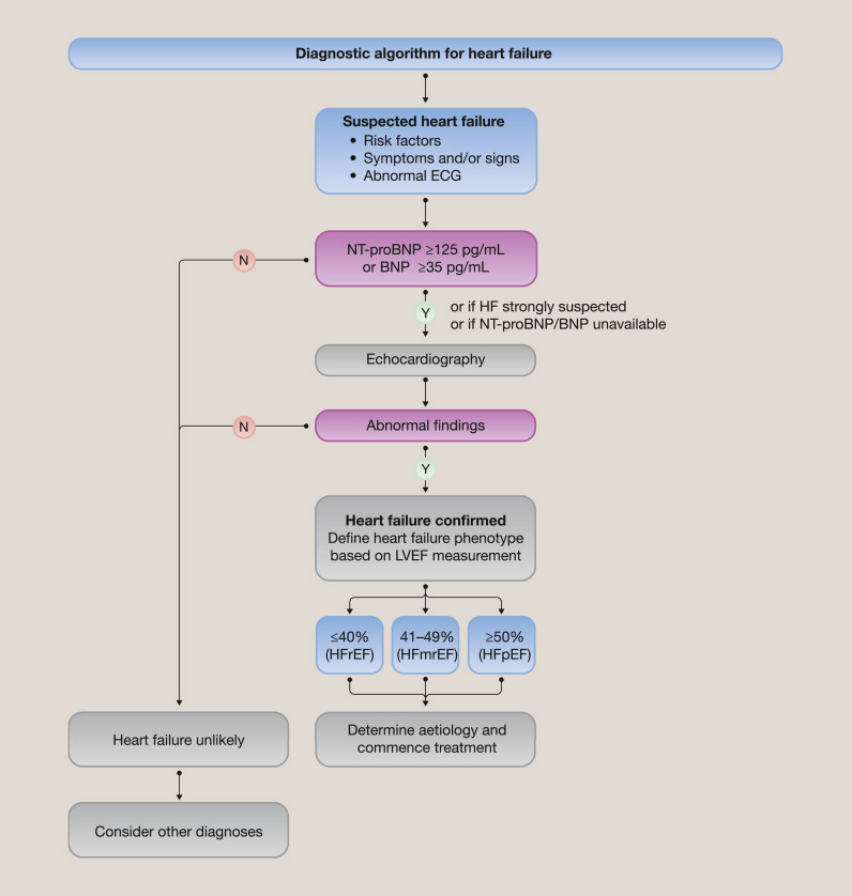

Figure 2.1: HF Diagnosis Pipeline from [8]

The guidelines provided by [8, 9] confirm or exclude the HF diagnosis on the basis of two tests: the measurement of the plasma concentration of natriuretic peptides (NPs) and transthoracic echocardiography (ECHO). The latter is used to extract indices associated with heart functionality and hemodynamics, which enables a deeper understanding of underlying causes. The extracted indices serve to calculate LVEF, a central measure of left ventricular systolic function. It is defined as the fraction of chamber volume ejected in systole (Stroke Volume (SV)) in relation to the volume of the blood in the ventricle at the end of diastole (End-Diastolic Volume (EDV)) Kosaraju et al. [10].

Clinical guidelines classify HF phenotypes depending on the measured LVEF:

Table 2.1: Types of Heart Failure and Associated Criteria

Type of Heart Failure (HF)	**Criteria**
HFrEF (Reduced Ejection Fraction)	LVEF $\leq 40\%$
HFmrEF (Mid-range Ejection Fraction)	$40\% <$ LVEF $\leq 50\%$
HFpEF (Preserved Ejection Fraction)	LVEF $> 50\%$ (and raised NPs confirm diagnosis)

In addition to these indices, experts in the field recognize the significant potential of Cardiac Time Intervals (CTIs) for assessing the condition of patients with HF [11].

CTIs offer several advantages, particularly for long-term monitoring and routine evaluations. Furthermore, they can also be derived from non-invasive signals such as ECG and PCG, making them simpler, more cost-effective, and less resource-intensive compared to ECHO, which requires specialized equipment and operator expertise [12] [13]

Unlike ECHO, these methods provide metrics for assessing ventricular performance with minimal inter-operator variability. Additionally, their accessibility and ease of repeated use make them ideal for monitoring disease progression and treatment effects over time.

Although ECHO remains essential for detailed anatomical and functional assessments, ECG/PCG-based CTIs complement them as a practical and efficient alternative in many clinical scenarios [10].

The following section examines how and which CTIs can be extracted from biomedical signals like PCG and ECG, and how these extracted intervals can discriminate LVEF intervals.

2.2. Cardiac Time Intervals

Cardiac Time Intervals (CTIs), also known as Systolic Time Intervals (STIs), are a measure of the durations between electrical and mechanical events within the cardiac cycle, most commonly measured in milliseconds. The relationship between the duration of the phases of left ventricular systole and the hemodynamics of the cardiac cicle allows for evaluation of the ventricular performance.

2.2.1 CTI's from Simultaneous PCG and ECG recording

Traditionally, CTIs are measured using echocardiography. However, current research focuses on developing alternative methods based on simultaneous recordings of biomedical signals that can be used to extract fiducial points that bound them. Phonocardiography (PCG), impedance cardiography (ICG), and more recently, seismocardiography (SCG) have been the methods of choice for detection of Cardiac Time Intervals, when synchronized with ECG. However, in healthy subjects, only SCG and PCG signals provided a acceptable accuracy and precision in estimating cardiac timings, as compared to ICG [14].

In this research, we will evaluate the CTI extracted from simultaneous recordings of ECG and PCG, collected with the Rijuven Cardiosleeve multimodal stethoscope*.

Figure 2.2 illustrates the relationship between the collected signals and the physiological events of the cardiac cycle. The remainder of this section focuses on understanding how these mechanical and electrical events are reflected in the signals, enabling their use as biomarkers for physiological processes.

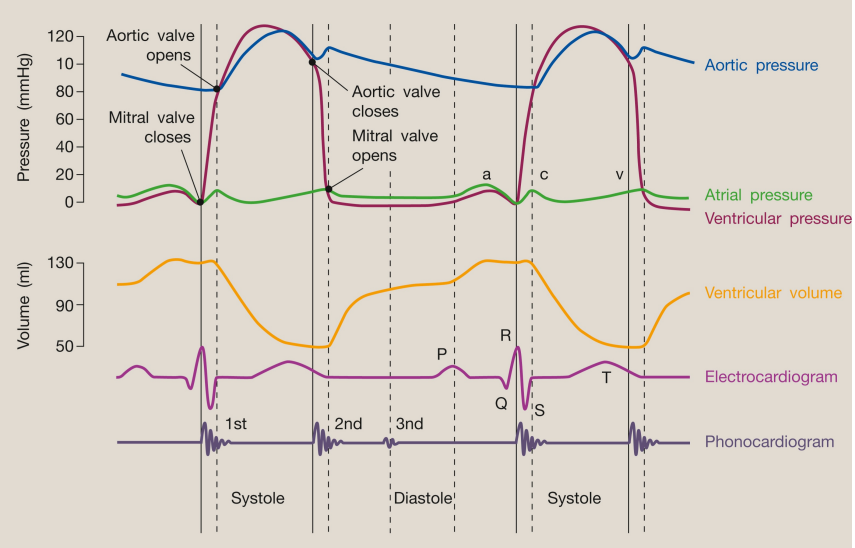

Figure 2.2: Wiggers diagram for the left heart function. Adapted from [15]

*https://www.rijuven.com/cardiosleeve

2.2.2 Phonocardiogram, PCG

PCG is the measure of the heart sounds and is captured using a stethoscope and microphone. These sounds are generated by valve closure as well as by blood flow turbulence during systole and diastole. In a normal heart, two dominant sounds, S1 and S2, appear in rhythmical form. These sounds are described in literature [3] has follows:

S1, the first heart sound Occurs at the onset of the ventricular systole.

Its first component is usually attributed to the closing of the atrio-ventricular (AV) valves (mitral and tricuspid), while its second high-frequency component is typically associated with the vibration induced by the left ventricle ejection.

S2, the second heart sound Occurs at the end of systole which is also the onset of the ventricular diastole.

It is composed of two main high-frequency components, which are highly correlated with the closing of the aortic and the pulmonary valves, respectively.

Spectral characteristics S1 and S2 are primarily within $10\,\mathrm{Hz}$–$300\,\mathrm{Hz}$, with focused energy in $50\,\mathrm{Hz}$–$150\,\mathrm{Hz}$ (S1) and $50\,\mathrm{Hz}$–$200\,\mathrm{Hz}$ (S2) [16, 17]. Murmurs can extend up to $600\,\mathrm{Hz}$, but core S1/S2 components rarely exceed $250\,\mathrm{Hz}$ [18].

Table 2.2: Key heart-sound frequency bands and sampling strategy

Parameter	Typical Range
S1 frequency	$50-150\,\mathrm{Hz}$
S2 frequency	$50-200\,\mathrm{Hz}$

Sampling recommendations Nyquist–Shannon theorem suggests sampling higher than 2× maximum signal frequency. For PCG, sampling at ≥1 kHz ensures faithful capture of S1/S2 and murmurs [17, 18]. Accordingly, many studies record at high sampling frequencies (e.g. $2\,\mathrm{kHz}$) and downsample to $1\,\mathrm{kHz}$ [19].

The obtained CardioSleeve ECG records are sampled at $3\,\mathrm{kHz}$ and then downsampling is performed to $1\,\mathrm{kHz}$, which meets Nyquist criteria, reduces data size, and retains diagnostic PCG content, aligning with best practices [16–18].

2.2.3 Eletrocardiogram, ECG

In this study, we analyze electrocardiographic signals acquired from a single-lead ECG. While single-lead ECGs offer limited spatial resolution compared to standard 12-lead ECGs, they are increasingly used in wearable and mobile health applications due to their portability, ease of use, and effectiveness in rhythm monitoring. Bellow is a brief outline the key differences between standard 12-lead ECGs and single-lead wearable ECG systems:

Comparison: 12-Lead ECG vs. Single-Lead Wearable ECG

- **12-lead ECG:**
  - Uses 10 electrodes to capture 12 leads, providing a comprehensive spatial (3D) view of cardiac electrical activity.
  - Gold standard for diagnosing ischemia, infarction, conduction disturbances, hypertrophy, etc.[20, 21].
- **Single-lead (wearable) ECG:**
  - Typically records lead I (wrist, finger, patch) and offers limited spatial information.
  - More prone to motion/artifact noise; used in ambulatory/home settings [22].

Fiducial points in specific wave peaks (P, Q, R, S, and T) The ECG waveform results from a series of coordinated electrical events across the conduction system of the heart, capturing the cumulative signals. These signals arise from action potential propagation across cardiac cells, resulting in changes in membrane potential [23].

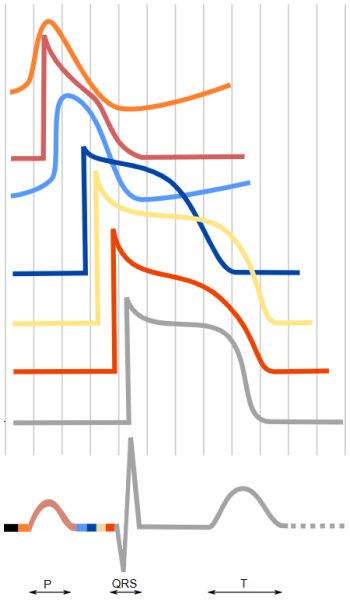

Figure 2.3: The superposition of all the heart action potentials resulting in the ECG signal.

P wave Represents the depolarization of both atria and the onset of atrial contraction.

The normal cardiac cycle begins with the firing of the sinoatrial node, which is not detected by a typical ECG. The depolarization of the sinoatrial node is then conducted rapidly throughout the right and left atria, giving rise to the P-wave. As the P-wave ends, the atria are thus depolarized and this relates to their contraction. The signal then returns to baseline while action potentials (not large enough to be detected) spread through the atrioventricular node and bundle of His.

QRS complex Represents ventricular depolarization and the onset of ventricular contraction.

Roughly 200 ms after the beginning of the P-wave, the right and left ventricles begin to depolarize resulting in the recordable QRS complex, which is approximately 100 ms in duration.

The first negative deflection (if present) is the Q-wave, the large positive deflection is the R-wave, and if there is a negative deflection after the R-wave, it is called the S-wave.

As the QRS complex ends, the ventricles are completely depolarized and are contracting. Importantly, the exact shape of the QRS complex depends on the placement of electrodes from which the signals are recorded. Simultaneous with the QRS complex, atrial contractions have ended and the atria are repolarizing.

T wave Represents ventricular repolarization.

Toward the end of ventricular contraction, the ECG signal returns to baseline. The ventricles repolarize after contraction, giving rise to the T-wave. The T-wave is normally the last detected potential in the cardiac cycle; thus, it is followed by the P-wave of the next cycle, repeating the process.

2.2.4 CTIs extraction from ECG and PCG

Since these two signals capture limited information of the heart's electrical and mechanical events, the start and end points of certain CTIs may either be directly present in the ECG and PCG signals or require inference. The following table summarizes the key CTIs described in the literature [14, 24] and indicates which intervals can be directly derived or inferred from these signals:

Table 2.3: Cardiac Time Intervals: PCG and ECG Analysis

Interval	Physiological Event	Fiducial Point	Present or Inferred
Electromechanical Activation Time (EMAT)	Onset of ventricular depolarization (Start)	ECG (Q-wave)	Present
	Mitral valve closure (End)	PCG (S1)	Present
Pre-Ejection Period (PEP)	Onset of ventricular depolarization (Start)	ECG (R-wave or Q-wave)	Present
	Aortic valve opening (End)	PCG (S1)	Inferred
Isovolumic Contraction Period (IVCP)	Mitral valve closure (Start)	PCG (S1)	Present
	Aortic valve opening (End)	PCG (S2)	Inferred
Left Ventricular Ejection Time (LVET)	Aortic valve opening (Start)	PCG (S1)	Inferred
	Aortic valve closure (End)	PCG (S2)	Present
Left Ventricular Systolic Time (LVST)	Mitral valve closure (Start)	PCG (S1)	Present
	Aortic valve closure (End)	PCG (S2)	Present
Total Systolic Time (TST) or Total Electromechanical Systole (QS2)	Onset of ventricular depolarization (Start)	ECG (R-wave or Q-wave)	Present
	Aortic valve closure (End)	PCG (S2)	Present

From Table 2.3, it is evident that not all physiological events correspond directly to those captured by Electrocardiogram (ECG) and Phonocardiogram (PCG) signals. Nevertheless, intervals such as Pre-Ejection Period (PEP), Left Ventricular Ejection Time (LVET),

Total Systolic Time (TST) and Electromechanical Activation Time (EMAT) can still be reliably estimated due to the close temporal proximity between the actual physiological events and the detectable signal features used for their calculation.

2.2.5 Derived measures

It should be noted that CTIs are inversely proportional with the duration of the cardiac cycle, as a result most researchers compute corrected CTIs by dividing the CTIs value by the RR interval. Another more popular approach is to calculate two relative dimensionless CTIs ratios, which have been defined as:

- **EMAT/ LVST**: The ratio between the electromechanical delay (EMAT) and the left ventricular systolic time (LVST), based on the closure of the mitral and aortic valves.
- **PEP/ LVET**: The ratio between the pre-ejection period (PEP) and the left ventricular ejection time (LVET), based on the opening and closure of the aortic valve.

The following section explores how these inferred intervals were found to be clinically feasible when manually extracted from ECG and PCG.

2.2.6 PCG/ECG manually annotated Cardiac Time Intervals's and Heart Failure

Figure 2.4 illustrates the relative positions of the intervals and their fiducial points:

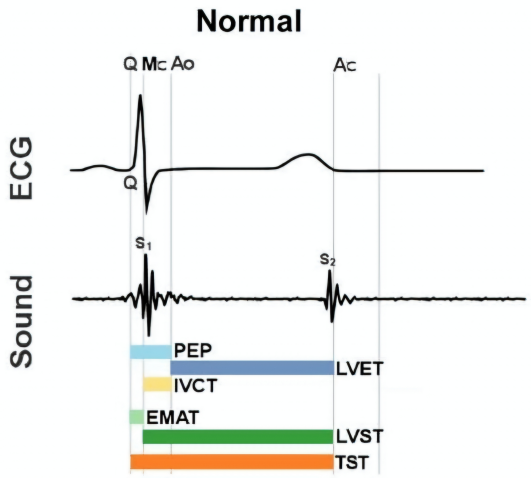

Figure 2.4: Cardiac Time Intervals delineated by PCG and ECG. Adapted from [25]

The clinical significance of manual annotations in PCG and ECG must be considered in order to determine whether an automatic approach can achieve comparable accuracy.

Numerous studies have been conducted to evaluate the feasibility of using simultaneous PCG and ECG recordings for estimating these intervals. The findings in the literature are as follows in Table 2.4:

Table 2.4: Measurement errors (*ME*) and correlation coefficients (*r*) of the PEP, LVET, TST, and EMD intervals estimated using different signals synchronous PCG and ECG according to the literature

Source	CTI	ME (ms)	r	Ref. data	Approach
Carvalho et al (2009)	**PEP**	5.8 ± 4.9	0.91	Echo	Manual, R wave
	LVET	14.8 ± 10.9	0.84		
Dehkordi et al (2019)	**TST**	1.22 ± 10.9	0.94	Echo	Manual, Q wave
	EMAT	5.3 ± 9.7	0.97	TDI	

Furthermore, literature [26, 27] highlights associations between CTIs extracted from ECG and PCG and factors such as LVEF, heart function, and clinical outcomes. They outlined that CTIs, such as EMAT, c EMAT, LVST, and PEP/ LVET, have shown significant promise as biomarkers for assessing heart failure. In [28], the following modifications were described to occur in the CTIs in cases of worsening cardiac functionality due to HF:

- c EMAT and c PEP increase.
- c LVST and c LVET decrease.
- EMAT/ LVST and PEP/ LVET increase.

Where *c* stands for corrected CTIs.

Another interesting interval is QT, which represents the time interval when the heart's ventricles depolarize and repolarize. A study [29] has found that QT duration (electrical systole) and QS2 duration (electromechanical systole) are closely aligned in healthy individuals, with QT being slightly shorter. It explored the higher QT and shorter QS2 was found to be a strong predictor of mortality (especially sudden death) after myocardial infarction.

Now that the relevance and feasibility of extracting these intervals from ECG and PCG has been established, we can proceed to analyze the automatic approaches for extracting these intervals as reported in the literature. The following chapter will provide a detailed review of the relevant studies.

3. Literature Review

The objective of this literature review is to identify methodologies for the automatic extraction of CTIs from simultaneous ECG and PCG recordings.

Given the importance of accurate and efficient CTIs measurement, this review focuses on exploring existing approaches that utilize ECG and PCG signals to automate CTIs extraction. This chapter aims to highlight key methodologies, their underlying principles, and to compare their performance. Furthermore, the review seeks to identify gaps and challenges in current research, providing a foundation for further investigation and potential advancements in this domain.

3.1. Research Questions

To guide the research performed in the scope of this thesis, the main research question to be investigated was carefully formulated:

Research Question: What methodologies have been developed to automatically extract CTIs from simultaneous ECG and PCG recordings?

3.2. Search Strategy

This section outlines the strategy and search queries used for the literature review, conducted in accordance with the PRISMA guidelines [30].

3.2.1 Databases

The literature search was performed using PubMed*, IEEE Xplore†, MDPI‡, and Scopus§, that were selected for their robust domain relevance and the breadth of relevant content.

3.2.2 Keywords

The keywords employed in this review were categorized according to their respective domains:

*https://pubmed.ncbi.nlm.nih.gov/
†https://ieeexplore.ieee.org/Xplore/home.jsp
‡https://www.mdpi.com/
§https://www.scopus.com/

- “ECG” OR ”EKG” OR “Electrocardiogram”
- “PCG” OR “Phonocardiogram”
- “CTI” OR “Cardiac Time Interval”
- “STI” OR “Systolic Time Interval”
- ”PEP” OR ”LVET” OR ”LVST” OR ”EMAT”

3.2.3 Queries

Search queries were designed and iteratively refined to identify the most relevant papers while minimizing irrelevant results. The process began with generic keywords and evolved to include more specific terms, ensuring better alignment with the research objectives.

The initial queries returned an overwhelming number of results, many of which were unrelated to the topic. By progressively adding targeted keywords and Boolean operators, the results were narrowed to achieve better results. Database-specific limitations on Boolean operators were also considered when structuring the queries.

The final queries and their differences are as follows:

Query 1: (”STI” OR ”Systolic Time Interval” OR ”CTI” OR ”Cardiac Time Interval”) AND (”ECG” OR ”EKG” OR ”PCG” OR ”Electrocardiogram” OR ”Phonocardiogram”)

Query 2: (”PEP” OR ”LVET” OR ”LVST” OR ”EMAT”) AND (”ECG” OR ”EKG” OR ”PCG”)

The difference between Query 1 and Query 2 lies in their specificity. Query 1 was broader, targeting general Cardiac Time Intervals and signal modalities, while Query 2 focused on specific cardiac parameters, refining the search to more precise and relevant results.

Table 3.1: Results of Query 1 and Query 2 across selected databases.

Database	Query 1 Results	Query 2 Results
PubMed	92	199
IEEE Xplore	35	54
MDPI	177	8
Scopus	232	30

3.2.4 Selection Process

The selection process for relevant papers was conducted in four stages:

1. Elimination of Duplicates: Identical records across databases were identified through the title and removed. Manual removal was necessary because of missing characters(such s ”.”) and lower and upper characters misalignment.
2. Article Characteristics: Records that did not aling with desired characteristics were not considered
3. Title Evaluation: Records were filtered based on relevance indicated by their titles.
4. Abstract Evaluation: Abstracts of shortlisted papers were reviewed to assess their alignment with the research objectives.
5. Full-Text Review: A through examination of the remaining papers was conducted to determine their relevance, including cross-referencing citations for additional pertinent studies.

Inclusion Criteria:

- **Article Characteristics:** Article published from 2014 onwards; Peer-reviewed journal/conference publication; Full-text availability; Written in English Language.
- **Title:** The title must explicitly indicate relevance to CTIs and (ECG or PCG).
- **Abstract:** The abstract should clearly describe automatic methodologies for extracting features from ECG and PCG.
- **Full-text Review:** Article extracts at least one cardiac time interval.

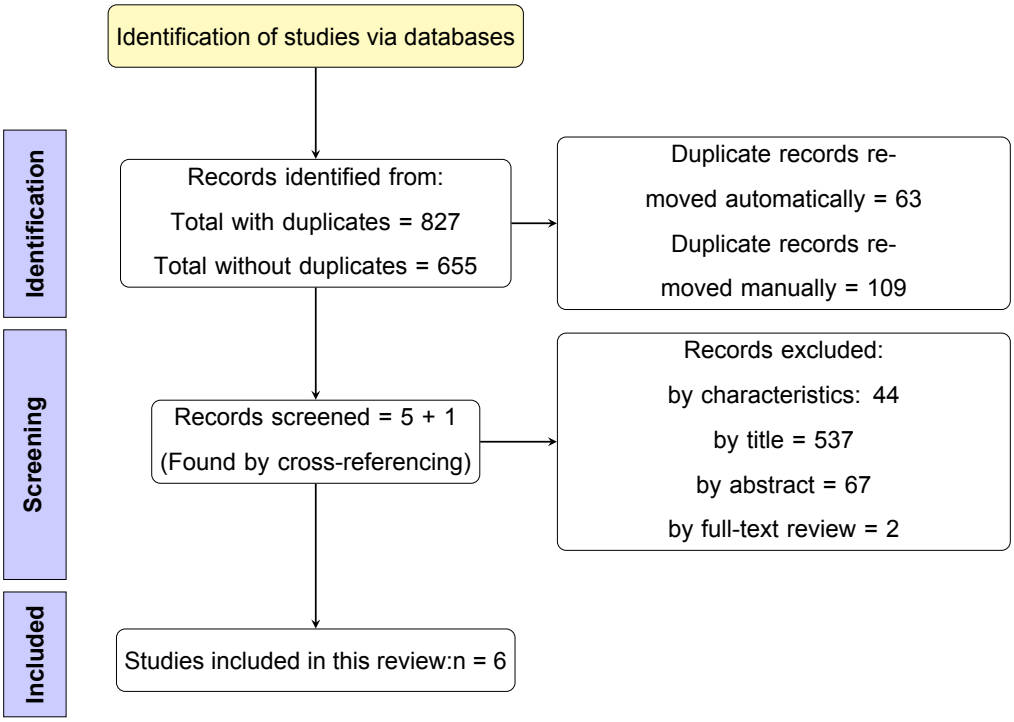

Figure 3.1: Paper screening flowchart, following PRISMA guidelines

3.3. Review on Automatic Extraction of CTIs from PCG and ECG

The final 6 articles indicate that two main approaches have been explored for detection of fiducial points. The first involves signal processing techniques, such as instantaneous energy, frequency analysis, and Bayesian or wavelet methods (Carvalho et al., 2010 [31]; Paiva et al., 2009 [32], 2012 [3]; Janjarasjitt, 2020 [33]). The second relies on identifying specific components and features of S1 and S2 (Dehkordi et al., 2019 [14]; Klum et al., 2020 [7]).

Performance results in table bellow, present measurement error (ME) and the correlation coefficient (*r*) obtained from these studies.

Table 3.2: Measurement errors (*ME*) and correlation coefficients (*r*) of the PEP and LVET intervals according to the literature. Adapted from [34]

Source	**PEP_{PCG}**		**$LVET_{PCG}$**		**Ref. data**	**Approach**
	ME (ms)	*r*	*ME (ms)*	*r*		
Carvalho et al. (2009) [35]	5.8 ± 4.9	0.91	14.8 ± 10.9	0.84	Echo	Manual, R wave
Paiva et al. (2009) [32]	10.3 ± 7.3	0.47	15.8 ± 13.6	0.77	Echo	Automatic, R wave
Carvalho et al. (2010) [31]	9.0 ± 6.4	0.54	14.4 ± 12.4	0.77	Echo	Automatic, P wave
Paiva et al. (2012) [3]	7.7 ± 5.9	0.51	11.4 ± 9.0	0.87	Echo	Automatic, Q wave
Janjarasjitt et al. (2020) [33]	13.4 ± 11.4	–	27.1 ± 33.4	–	Echo	Automatic, R wave
Klum et al. (2020) [7]	17.6 ± 17.9	–	21.8 ± 23.78	–	ICG	Automatic, Q wave
Jiménez et al. (2021) [34]	H^1: -12.8 ± 13.4	0.5	33.4 ± 27.2	0.7	Echo	Automatic, R wave
	C^1: -16.3 ± 13.8	0.0	42.9 ± 26.9	0.6		
	H^2 -6.6 ± 19.6	0.0	24.4 ± 32.4	0.7		
	C^2: -22.8 ± 24.7	0.0	44.6 ± 33.6	0.6		
	H^3: 0.3 ± 8.3	0.8	24.4 ± 32.4	0.7		
	C^3: -0.2 ± 7.5	0.8	44.6 ± 33.6	0.6		

By H^1 and C^1 the peak of the envelopes of S1 and S2, H^2 and C^2 the local maximum of S1 and S2, by H^3 and C^3 a point at about the peak of S1 and S2. H and C respectively stand for healthy and cardiac disease groups

3.3.1 Carvalho et al., 2010 [31]; Paiva et al., 2009 [32], 2012 [3]

Carvalho et al., 2010 [31] and Paiva et al., 2009 [32], 2012 [3] investigate PEP and LVET, through Bayesian signal processing approaches, validated by comparing results with echocardiography, the clinical gold standard. In addition to slight changes in the algorithm approach, they also differ in the fiducial points used to estimate the beginning of PEP as described in the 'Approach' column of table 3.2.

The 2009 and 2010 studies focused only on healthy individuals, utilizing the instantaneous amplitude (IA) of the heart sound waveform derived via the Hilbert Transform. PEP was calculated as the interval between the ECG R-peak or P-wave and aortic valve opening, which was identified as valleys in the IA curve, assumed to occur 30 ms after atrioventricular (AV) valve closure. LVET was determined using aortic valve closure detected via Fast Wavelet Transform (FWT) and Shannon energy analysis.

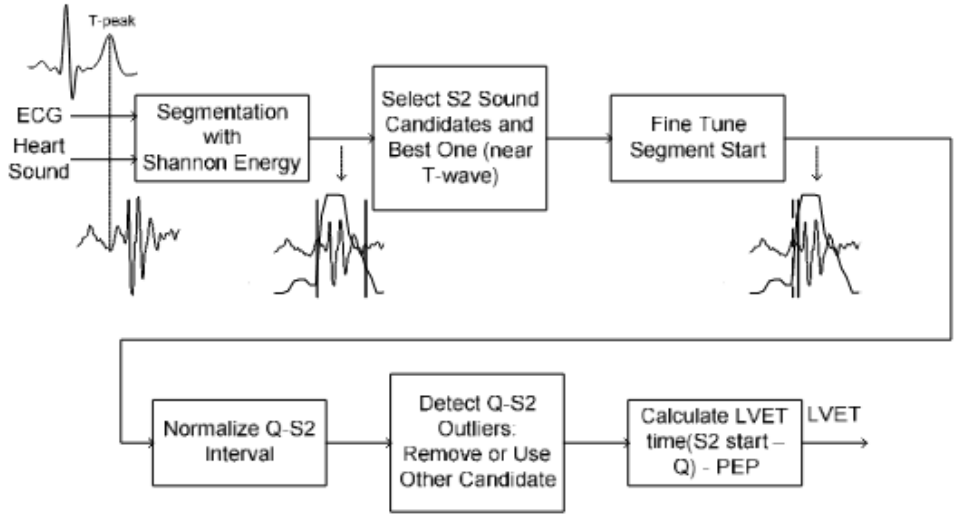

Figure 3.2: LVET extraction flow-chart. Obtained from [3]

Building on this, the 2012 study applied similar methodologies to both healthy subjects and patients with cardiovascular diseases (CVD). The Bayesian approach incorporated a two-pass algorithm to refine estimates using subject-dependent Gaussian models. The two-pass approach yielded noticeable improvements in PEP estimation accuracy compared to a single-pass model.

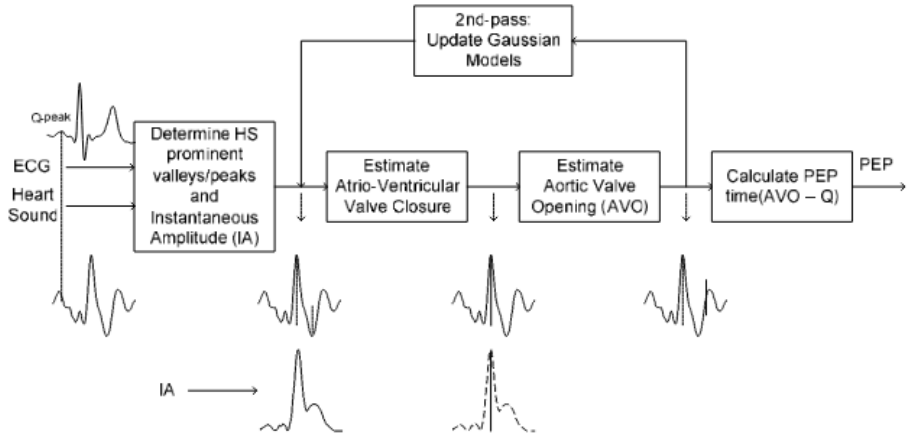

Figure 3.3: PEP extraction flow-chart. Obtained from [3]

3.3.2 Janjarasjitt et al., 2020 [33]

This study presents an approach based on continuous wavelet transform (CWT) of synchronized ECG and PCG signals. The method was validated using data collected from 68 volunteers, split into healthy and cardiovascular disease (CVD) groups.

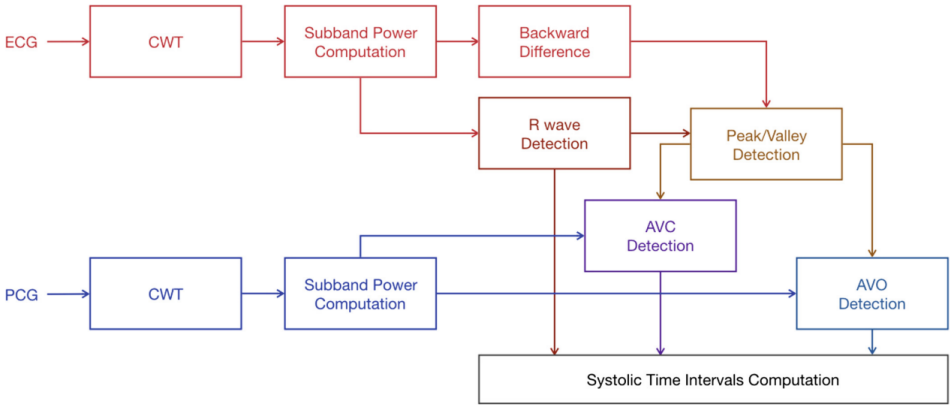

Figure 3.4: PEP and LVET extraction pipeline. Obtained from [33]

The processing of ECG and PCG signals involves distinct steps to extract meaningful features. For the ECG signal, continuous wavelet transform (CWT) is applied, followed by the analysis of subband power, detection of R-wave peaks, and the calculation of derivatives to identify key characteristics. Similarly, for the PCG signal, CWT is used to analyze subband power and determine the timing of aortic valve events, such as opening and closure, based on signal peaks and valleys. These processes aim to extract relevant information for further analysis.

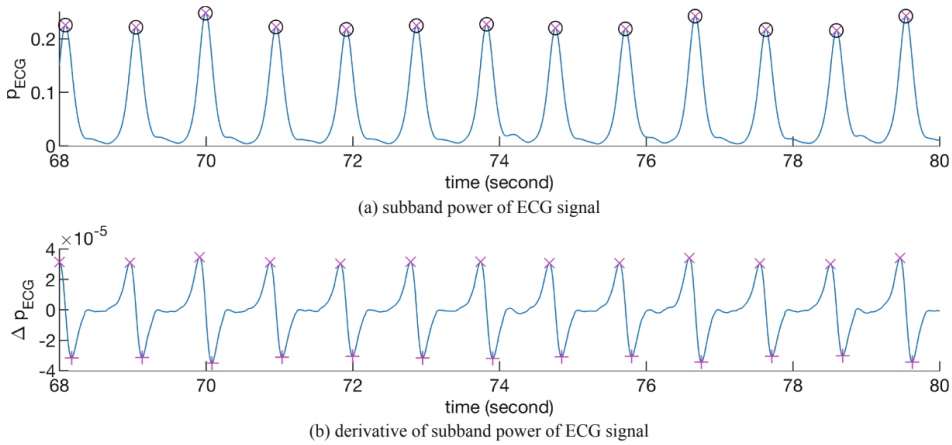

Figure 3.5: Extracted peaks and valleys of exemplary ECG (a) subband power and (b) derivative of subband power. Obtained from [33]

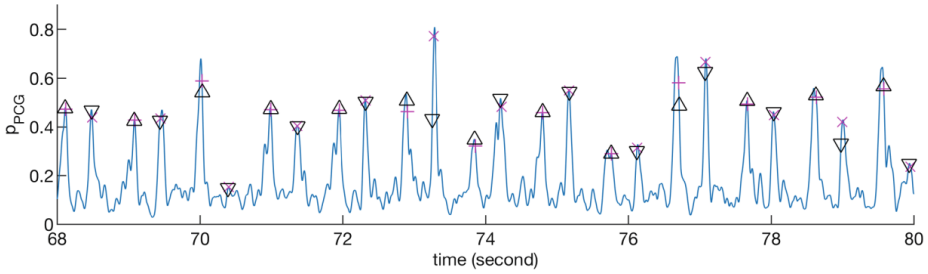

Figure 3.6: Extracted peaks of exemplary PCG subband power. Obtained from [33]

In the final stage, PEP and LVET are calculated using the timings of the R-wave and the aortic valve events. Specifically:

$$PEP = t_{S1\_peak} - t_R,$$

$$LVET = t_{S2\_peak} - t_{S1\_peak},$$

3.3.3 Klum et al., 2020 [7]

This study investigates the use of key features such as heart rate, QRS amplitude, and heart sounds (S1 and S2) to estimate PEP, LVET, and respiratory parameters.

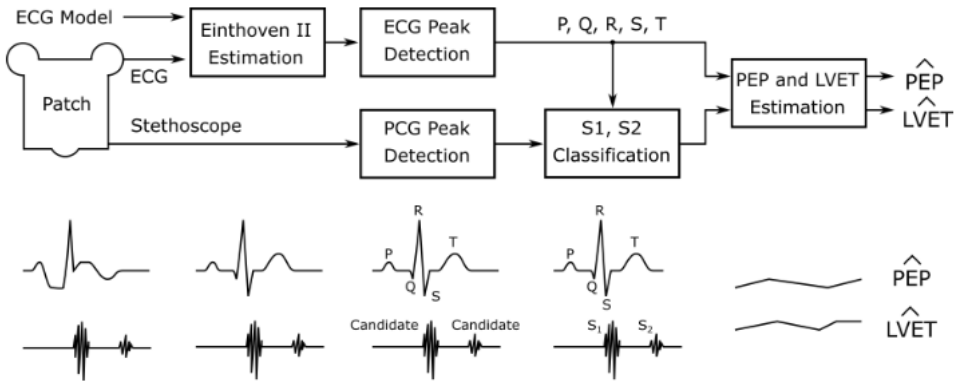

Figure 3.7: Ilustration of the algorithm created for PEP and LVET extraction. Adapted from [7].

Model Classes and Regression Analysis The authors used three model classes for regression analysis: polynomials (up to the third order), multi-layer perceptrons (MLPs), and time-delay neural networks (TDNNs). Using the three models, the patch ECG signals were fitted to estimate Einthoven Lead I and to estimate Einthoven Lead II. These models were trained on normalized data for each subject and position. Model performance was optimized based on the minimum Bayesian Information Criterion (BIC).

PEP and LVET Estimation ECG signals were pre-processed using resampling, low-pass filtering, and baseline removal. PCG signals were band-pass filtered and resampled. The Pan-Tompkins algorithm was used for R-peak detection in ECG, and peak detection for candidate S1 and S2 in PCG was carried out using an empirical wavelet transform, later on classified with a scheme of set rules. PEP and LVET were derived as the time differences between specific peaks in ECG and PCG:

$$PEP = t_{S1\_peak} - t_{Q},$$
$$LVET = t_{S2\_peak} - t_{S1\_peak},$$

PEP and LVET Estimation from ECG and PCG Peak detection was more accurate in ECG signals compared to PCG, with challenges arising from noise and breathing-related artifacts.

Overall, the combined use of these signals enhanced the accuracy of parameter estimation, despite challenges from noise and signal quality.

3.3.4 Jiménez-Gonzalez, 2021 [34]

This study utilized Semi-blind Canonical Independent Component Analysis (SCICA) to process ECG and PCG signals for the automatic detection of the Pre-Ejection Period and Left Ventricular Ejection Time. The dataset consisted of 36 recordings with annotated R peaks, aortic valve opening (AVO), and aortic valve closure (AVC), as well as reference PEP and LVET intervals. Noise sources, such as coughing and talking, posed challenges that were addressed through a three-stage signal processing pipeline.

In the first stage, SCICA was used to denoise the PCG signals by extracting independent components (ICs) and selecting cardiac-related ICs based on their frequency patterns and autocorrelograms. This improved the signal-to-noise ratio (SNR) and resolved morphological issues, enabling clearer identification of heart sounds S1 and S2.

The second stage focused on detecting AVO and AVC by extracting peaks (pS1, pS2) from the smoothed PCG envelope, generated using the Hilbert transform and band-pass filtering. Candidate peaks were identified near R peaks and validated using interval comparisons and adjacent heartbeat data. The detection of turning points (tS1, tS2) provided more stable markers for timing measurements.

The final stage evaluated performance by comparing p-based (pS1, pS2) and t-based (tS1, tS2) intervals with echocardiogram (Echo) annotations. T-based intervals consistently yielded smaller errors and better correlation with reference timings. P-based timings were more affected by noise and exhibited systematic underestimation of PEP and overestimation of LVET.

3.3.5 Dicussion on Research

Despite numerous advances in ECG–PCG interval estimation, several persistent challenges remain.

First, noise sensitivity remains a critical issue: heuristic peak-picking on single cardiac cycles is easily disrupted by transient artifacts—such as coughs or muscle noise—often yielding spurious PEP and LVET values.

Second, transform-domain approaches (e.g. wavelet or Hilbert analyses) demand meticulous parameter tuning—selecting appropriate scales, bandwidths, and detection thresholds—typically through manual or semi-empirical trial-and-error.

Finally, the use of fixed filters limits adaptivity: predefined frequency bands and filter shapes cannot always accommodate the wide variety of patient specific heart sound morphologies (for example, low-amplitude S2 sounds or pathological murmurs), reducing the reliability of event detection across diverse clinical populations.

Building on established ECG–PCG synchronization and transform-domain methods (Paiva et al. 2009; Janjarasjitt 2020; Jiménez-González 2021) and recent machine-learning approaches to PEP/ LVET estimation (Klum et al. 2020), our pipeline tackles three key challenges.

First, we form a median heartbeat template by aligning individual beats at their R-peaks and taking the median across cycles, which attenuates transient artifacts (e.g., coughs or muscle noise) so that peak-picking operates on a denoised representative beat rather than any single noisy cycle.

Second, a trained U-Net Neural Network is used to delineate the precise event time windows, thereby learning the optimal features and temporal scales directly from data and eliminating manual tuning of wavelet or Hilbert-transform parameters.

Finally, because the U-Net adapts its filters during training, it naturally accommodates patient-specific morphologies—including low-amplitude sounds and murmurs—without relying on fixed, one-size-fits-all filters. Within each segmented window, We then identify the event start, end, and center points, which serve as the onset and offset markers for computing a composite Cardiac Time Intervals, yielding a robust, adaptive framework that addresses noise sensitivity, parameter-tuning burden, and limited generalizability.

In the following chapter, each step of the CTI-extraction framework is described in detail.

4. Materials

This chapter provides a detailed description of the dataset used in this study.

4.1. ULSGE Dataset: Data Collection and Population

Data Collection The dataset used in this study was collected in collaboration with the Cardiac Department of the Unidade Local de Saúde Gaia e Espinho (ULSGE). This research is part of the Health from Portugal (HfPT) National Project.

Synchronous Electrocardiogram and Phonocardiogram recordings were acquired using the FDA-approved Rijuven Cardiosleeve multimodal stethoscope. The PCG signals were initially sampled at 3000 Hz while the ECG signals were sampled at 500 Hz.

Table 4.1: Age distribution of participants in the ULSGE dataset.

Age Group	Age Range	Percentage (%)
Children and Teenagers	≤ 19 years	< 1
Young Adults	20–34 years	< 1
Middle-Aged Adults	35–49 years	6
Older Adults	50–64 years	18
Seniors	≥ 65 years	74
Total	2–93 years	100

Records were obtained from 190 unique individuals, of which 167 had valid data. The population is approximately 60% males and 40% females, with ages ranging from 2 to 93 years. An age-stratified breakdown is shown in table 4.1.

Recordings were performed at four standard auscultation sites: the aortic, pulmonary, tricuspid, and mitral valves, each lasting approximately 30 seconds. In total, 558 pairs of synchronous ECG and PCG recordings were collected.

Annotations Only a few patients were categorized based on their Left Ventricular Ejection Fraction using the following three groups:

- **Group 1 (Reduced Ejection Fraction, rEF):** EF $< 40\%$
- **Group 2 (Mildly Reduced Ejection Fraction, mrEF):** EF 40–49%
- **Group 3 (Preserved Ejection Fraction, pEF):** EF $\geq 50\%$

A total of 160 recordings have annotated LVEF groups.

Manual annotations were taken from the median hearbeat template of each recording as reported in [36], accepted at Computing in Cardiology 2025. They consist of QRS onset, T-wave onset, First Heart Sound (S1), and Second Heart Sound (S2) fiducial points. From these, the following CTIs were derived: QT, QS2, S1S2. Due to signal-quality constraints, only recordings with well defined fiducial point onsets and offsets were annotated. Consequently, 84 recordings were selected for annotation. Of these:

- 4 recordings belong to patients with rEF,
- 26 recordings belong to patients with mrEF,
- 54 recordings belong to patients with pEF.

5. Methodology

In this research the objective is develop two methods of extracting Cardiac Time Intervals and performing a comparative analysis. Firstly, we employ a U-Net model to segment the signals Beat-by-Beat, for which we propose a strategy to post-process the outputed probabilities. Secondly, the other approach relies on pre-processing of the signals into median heartbeat templates, which will later be segmented by the model allowing us to obtain the CTIs. The flowchart bellow illustrates two methods' workflow.

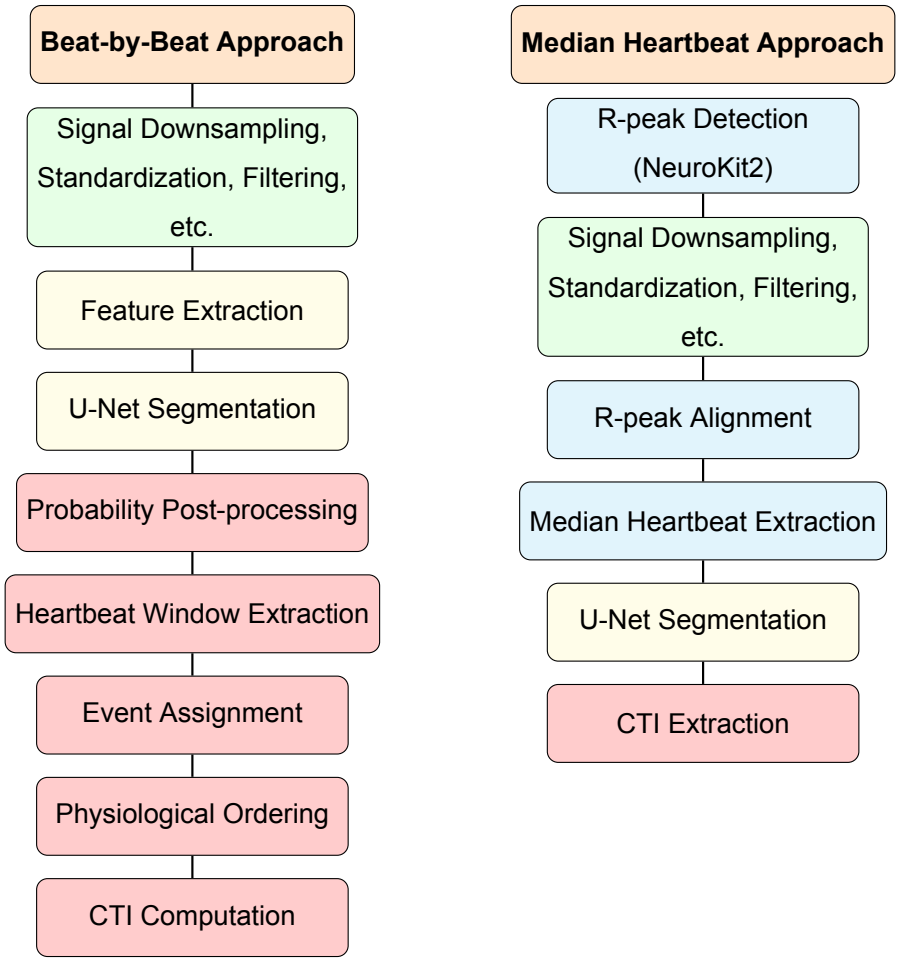

Figure 5.1: Overview of the segmentation pipelines: the left branch shows the beat-by-beat U-Net-based approach, while the right branch illustrates the median heartbeat approach. Steps common to both approaches are depicted in the same color

5.1. Signal Processing: Prepping signals for further computations

Prior to implementing either approach, ECG and PCG signals are subjected to a uniform preprocessing pipeline. In what follows, we describe each stage in detail and present the corresponding mathematical formulations.

1. Dataset Standardization The original dataset, represented by a table $\mathcal{D}$ with columns

$$\{\text{ID},\ \text{Auscultation Point},\ \text{ECG Signal},\ \text{PCG Signal}\}.$$

Each row i of $\mathcal{D}$ contains:

$$\text{ID}_i \in \mathbb{N}, \quad \text{Auscultation Point}_i \in \{\text{“Aortic”, “Pulmonic”, } \ldots\},$$
$$\text{Signal}_i = \Big[x_i[n]\Big]_{n=0}^{N_i-1},$$

where N_i denotes the number of recordings of subject i, respectively.

1. **Z-Score Standardization.** After, each retained ECG and PCG signal is independently standardized via

$$x_i[n] \longmapsto \widehat{x}_i[n] = \frac{x_i[n] - \mu_{x,i}}{\sigma_{x,i}},$$

This operation ensures $\mathrm{mean}(\widehat{x}_i) = 0$ and $\mathrm{std}(\widehat{x}_i) = 1$ for each signal.

Upon completion of this procedure, all subsequent operations reference the standardized dataframe $\mathcal{D}_{\text{std}}$.

2. Signal Filtering Once $\mathcal{D}_{\text{std}}$ is obtained, we process each subject's ECG and PCG independently. Denote the i-th subject's standardized signals by

$$\widehat{e}_i[n] \quad (\text{ECG}), \qquad \widehat{p}_i[n] \quad (\text{PCG}).$$

Let F_{ECG} be the ECG sampling frequency ($F_{\text{ECG}} = 500\,\text{Hz}$) and F_{PCG} be the PCG sampling frequency ($F_{\text{PCG}} = 3000\,\text{Hz}$).

2.1. ECG Preprocessing For each valid ECG signal $\widehat{e}_i[n]$, the following steps are applied in sequence:

1. **Notch Filter at 50 Hz.** To remove power-line interference at $50\,\text{Hz}$, an IIR notch filter with quality factor $Q = 30$ is applied. Afterwards, we apply zero-phase filtering via a forward-backward pass:

$$y_i[n] = \mathrm{filtfilt}\Big(\{b_k\}, \{a_k\}, \widehat{x}_i[n]\Big).$$

2. **Bandpass Butterworth (0.5–100 Hz).** Next, the recording is passed through a 6th-order Butterworth bandpass filter with cutoff frequencies $f_{\text{low}} = 0.5\,\text{Hz}$ and $f_{\text{high}} = 100\,\text{Hz}$.

3. **Median Removal.** Finally, any residual DC offset is removed by subtracting the sample median:
$$z_i[n] = y_i[n] - \text{median}\Big\{ y_i[k] \Big\}_{k=0}^{N_i-1}.$$

From theses steps results the fully processed ECG signal $\widetilde{e}_i[n]$.

2.2. PCG Preprocessing For each valid PCG signal $\widehat{p}_i[n]$, we perform the following:

1. **Downsampling** PCG signals are originally sampled at $f_{\text{orig}} = 3000\,\text{Hz}$. These signals are downsampled to $f_{\text{target}} = 1000\,\text{Hz}$. Downsampling uses a polyphase finite impulse response (FIR) filter. This process minimizes aliasing by applying a low-pass filter in conjunction with reducing the sampling rate, ensuring computational efficiency.

2. **Despiking via Schmidt Method.** To mitigate the impact of transient, a despiking algorithm described by Schmidt was implemented

After filtering, the final outputs for subject i are:

$$\widetilde{e}_i[n], \quad (n = 0, \ldots, N_i - 1), \qquad p_i[n], \quad (n = 0, \ldots, N_i - 1),$$

where N_i is the length of the processed signals.

5.2. Beat-by-Beat Approach

Research led by the Multiscope's invetigation group provided a U-Net model, conceptualized and trained by [37], accepted at Computing in Cardiology 2025, and its architecture was based on [38]. Two separate models were trained—one for ECG and one for PCG—both sharing the same U-Net architecture but using different features and label sets. A description of the model architecture, training and post-processing algorithm is done in the following sections.

5.2.1 Label Sets

Each model outputs four class-probability maps (four channels). The label sets are defined as follows:

- **PCG labels:**

$$\{0 : \text{S1},\ 1 : \text{systole},\ 2 : \text{S2},\ 3 : \text{diastole}\}.$$

- **ECG labels:**

$$\{0 : \text{baseline},\ 1 : \text{P-wave},\ 2 : \text{QRS},\ 3 : \text{T-wave}\}.$$

5.2.2 Preprocessing and Feature Extraction

Before the signals are input to the U-Net, both ECG and PCG signals undergo the preprocessing steps mentioned in 5.1 and feature extraction to enhance relevant characteristics and reduce noise.

ECG The ECG signal is transformed into a set of four time-domain features (or ”envelopes”):

1. **Hilbert Envelope:** Captures the amplitude variations by computing the analytic signal using the Hilbert transform.
2. **Shannon Energy Envelope:** Emphasizes high-energy segments by applying a logarithmic energy-based transformation.
3. **Homomorphic Envelope:** Highlights the signal envelope by computing the logarithm of the absolute signal values, applying a median filter (window size: 21 samples), and exponentiating the result.
4. **Smoothed Envelope:** Applies a smoothing operation using a 21-sample Hamming window to reduce high-frequency fluctuations.

PCG The PCG signal is downsampled to 1000 Hz and then undergoes a similar process, but the extracted features include both time and frequency domain information:

1. **Homomorphic Envelope:** Highlights the envelope of the PCG signal using a median filter, similar to the approach used for the ECG.

2. **Wavelet Scalogram (Morlet):** A time-frequency representation obtained using a continuous wavelet transform with the Morlet wavelet, focused on the frequency range of 20–200 Hz.

3. **Wavelet Scalogram (Mexican Hat):** Another time-frequency representation using the Mexican Hat wavelet, also covering the 20–200 Hz frequency range.

4. **Hilbert Envelope:** Captures the amplitude envelope of the signal using the Hilbert transform.

For both ECG and PCG, all four feature sequences are downsampled to 50 Hz to reduce the computational complexity of the model. They are then stacked to form a $4 \times L'$ matrix and split into patches of 64 samples with an 8-sample stride.

Finally, the ECG and PCG feature patches are separately fed into their respective U-Net models for segmentation.

5.2.3 Architecture

The segmentation strategy utilizes the U-Net architecture adapted from [38]. Both the ECG and PCG models share this same encoder–decoder topology:

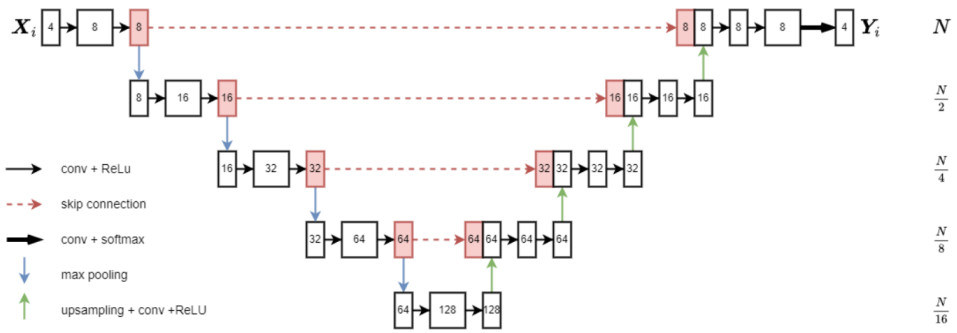

Figure 5.2: U-Net architecture used for heartbeat segmentation. Numbers inside boxes denote feature-map counts; dimensions on the right denote temporal length (samples) at each stage. Adapted from [38].

- **Input:** A 4×64 patch of feature envelopes.

- **Encoder:** Four downsampling blocks, each consisting of two 1D convolutions (kernel size 3), batch normalization, ReLU activation, and a 2× max-pooling. The number of feature maps doubles at each down-step (e.g., 16 → 32 → 64 → 128).

- **Bottleneck:** Two 1D convolutions with 128 feature maps, batch normalization, and ReLU.
- **Decoder:** Four upsampling blocks, each with a transposed convolution (stride 2), concatenation with the corresponding encoder feature map (skip connection), followed by two 1D convolutions (kernel size 3), batch normalization, and ReLU. Feature maps are halved at each up-step ($256 \rightarrow 128 \rightarrow 64 \rightarrow 32 \rightarrow 16$).
- **Output:** A final 1D convolution (kernel size 1) producing four channels, passed through softmax to obtain class probabilities for each sample in the 64-sample patch.

This architecture is particularly suited for capturing multi-scale temporal features while preserving precise localization via skip connections. Training was performed separately:

- **ECG model:** Trained on the LUDB dataset [39] with ECG patches and labels {*baseline*, *P-wave*, *QRS*, *T-wave*}. The LUDB comprises 200 subjects, each providing a single 10-second, 12-lead ECG recording (leads I, II, III, aVR, aVL, aVF, V1–V6) sampled at 500 Hz.
- **PCG model:** Trained on the PhysioNet 2016 dataset [40] with PCG patches and labels {*S1*, *systole*, *S2*, *diastole*}. This challenge set includes 3,153 training recordings from approximately 1,072 subjects, of which 2,302 are clean-normal, 186 noisy-normal, 572 clean-abnormal and 93 noisy-abnormal, all resampled to 2,000 Hz.

5.2.4 Post-Processing of U-Net Probabilities

The U-Net outputs a probability map indicating the likelihood of each sample belonging to each class. These raw probability outputs require a sequence of algorithmic steps to extract the cardiac intervals. This section describes the systematic workflow for post-processing U-Net segmentation probabilities, including extraction of events' start, center and middle points; conversion to time domain; heartbeat delineation; event to heartbeat association and interval computation. The sequence of steps in the pipeline is depicted in the flowchart bellow:

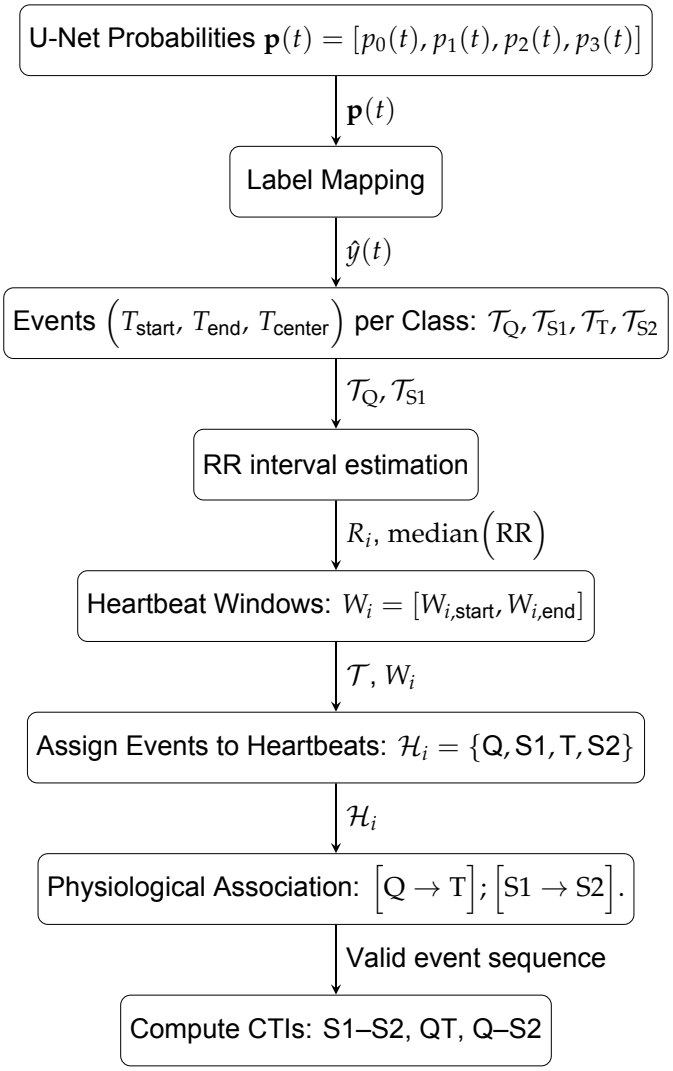

Figure 5.3: CTI extraction pipeline from U-Net output probabilities.

1. **Label Mapping and Event Extraction**: Convert per sample class probability vectors, outputed by the U-Net models, into discrete labels:

$$\mathbf{p}(t) = [p_0(t),\ p_1(t),\ p_2(t),\ p_3(t)],$$

The chosen approach was to apply an $\arg\max$ operation at each sample index t:

$$\hat{y}(t) \ = \ \underset{c \in \{0,1,2,3\}}{\arg\max}\ p_c(t)\,.$$

This yields a sequence of integer labels $\hat{y}(1), \hat{y}(2), \ldots, \hat{y}(N)$ for a recording of N time samples. Then, it is possible to identify the samples of the contiguous segments corresponding to target event classes. From each recording we obtain four lists of start, center and end times of the events' segments where each event is represented

as a triple

$$\left(T_{\text{start}},\, T_{\text{end}},\, T_{\text{center}}\right),$$

where

$$T_{\text{start}} = \frac{t_{\text{start}}}{f_s}, \quad T_{\text{end}} = \frac{t_{\text{end}}}{f_s}, \quad T_{\text{center}} = \frac{T_{\text{start}} + T_{\text{end}}}{2}$$

and for each recording we obtain time-based dictionaries such as

$$\mathcal{T} = \begin{cases} \text{Q} \;\mapsto\; \left[\left(T^{(QRS)}_{\text{start},i},\, T^{(Q)}_{\text{end},i},\, T^{(Q)}_{\text{center},i}\right)\right]_{i=1}^{n_Q}, \\ \text{T} \;\mapsto\; \left[\left(T^{(T)}_{\text{start},j},\, T^{(T)}_{\text{end},j},\, T^{(T)}_{\text{center},j}\right)\right]_{j=1}^{n_T}, \\ \text{S1} \;\mapsto\; \left[\left(T^{(S1)}_{\text{start},k},\, T^{(S1)}_{\text{end},k},\, T^{(S1)}_{\text{center},k}\right)\right]_{k=1}^{n_{S1}}, \\ \text{S2} \;\mapsto\; \left[\left(T^{(S2)}_{\text{start},\ell},\, T^{(S2)}_{\text{end},\ell},\, T^{(S2)}_{\text{center},\ell}\right)\right]_{\ell=1}^{n_{S2}}. \end{cases}$$

With the extracted timing we can now proceed to heartbeat delineation.

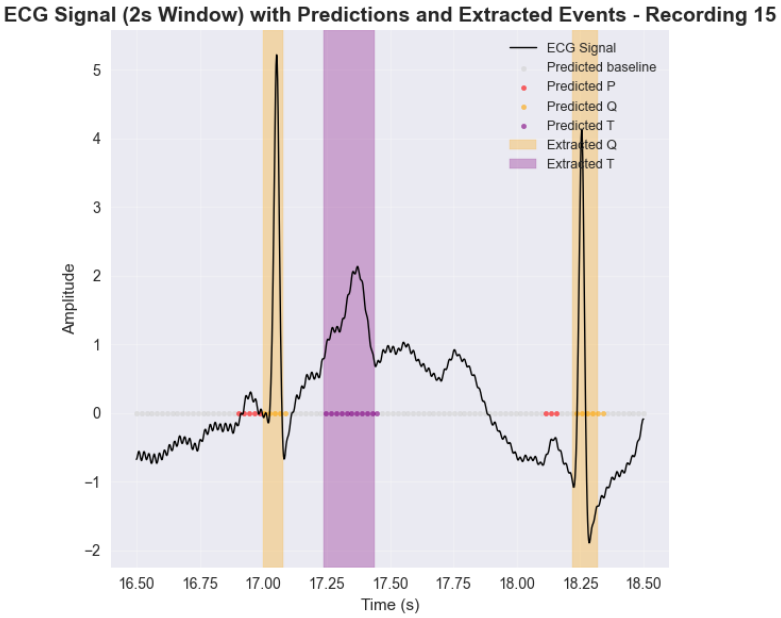

Figure 5.4: Segment of an ECG signal and the extracted windows for the events Q and T, returned by the U-Net model. From the Beat-by-Beat approach

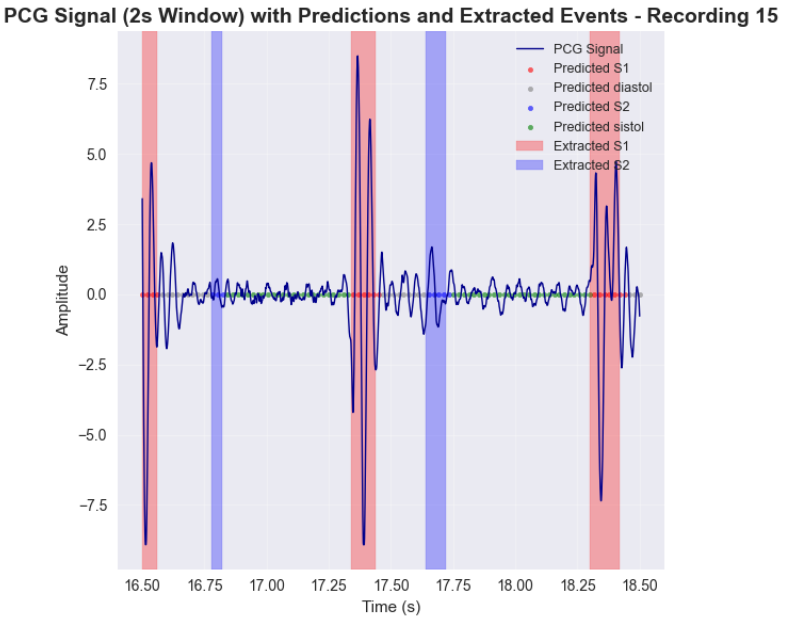

Figure 5.5: Segment of an PCG signal and the extracted windows for the events S1 and S2, returned by the U-Net model. From the Beat-by-Beat approach

2. **Heartbeat Window Detection**: The next goal is to identify discrete heartbeat windows. Ideally, the R-peaks in the ECG would serve as references for individual beats. In the Median Heartbeat approach, we will be able to use the Neurokit2 estimates of the R peaks. However, because the U-Net returns the QRS segments, we use the center of the segments as our estimate of the R peak in the Beat-By-Beat approach. Therefore, the algorithm follows a hierarchical approach to determine reference times:

   Firstly, we estimate the *median* RR interval empirically by taking differences between successive reference times:

   $$\Delta R_i = R_i - R_{i-1}, \quad i = 2, \ldots, M,$$

   and computing the median

   $$\text{median}\Big(\text{RR}\Big) = \text{median}\Big(\{\Delta R_i\}_{i=2}^{M}\Big).$$

   If only a single R-peak/QRS segment is found, we set RR interval value as the typical 0.8s.

   Secondly, in order to center the hearbeat window we use either R-peak (for the median heartbeat approach) or the QRS center times (for the beat-by-beat approach) as the center reference for beats.

Once a sorted list of reference centers $\{R_1, R_2, \ldots, R_M\}$ are established (where $R_1 < R_2 < \cdots < R_M$), heartbeat windows are defined as follows:

- For R_i, set

$$W_{i,\text{start}} = R_i - 0.5\,\text{median}\Big(\text{RR}\Big), \quad W_{i,\text{end}} = R_i + 0.5\,\text{median}\Big(\text{RR}\Big).$$

In this way, each reference time yields a heartbeat window

$$W_i = \Big(W_{i,\text{start}},\, W_{i,\text{end}}\Big).$$

The fraction 0.5 (50%) is an heuristic choice that account for the fact that the Q-wave typically occurs in the middle of systole and diastole.

Assignment of Events to Heartbeats: After defining heartbeat windows $\{W_i\}_{i=1}^{M}$, the next operation is to assign each event to exactly one heartbeat window. The assignment procedure proceeds as follows:

- *Center-Based Candidate Selection*

  For each event, check whether its center time T_{center} falls within any window $W_i = \Big(W_{i,\text{start}},\, W_{i,\text{end}}\Big)$, i.e.,

$$W_{i,\text{start}} \;\leq\; T_{\text{center}} \;\leq\; W_{i,\text{end}}.$$

  If multiple windows satisfy the center this criterion (which can occur at boundaries), compute for each candidate window the amount of overlap between the event's time span $\Big[T_{\text{start}},\, T_{\text{end}}\Big]$ and the window's span:

$$\text{Overlap}_i = \min\Big(T_{\text{end}},\, W_{i,\text{end}}\Big) \;-\; \max\Big(T_{\text{start}},\, W_{i,\text{start}}\Big),$$

  and select the window with the maximum positive overlap. If exactly one window contains the event center, assign the event to that window immediately.

- *Fallback Overlap Search*

  If no window contains T_{center}, check all windows for nonzero overlap in the sense

$$\text{Overlap}_i = \min\Big(T_{\text{end}},\, W_{i,\text{end}}\Big) - \max\Big(T_{\text{start}},\, W_{i,\text{start}}\Big) > 0.$$

  Assign the event to whichever window yields the maximum overlap. This ensures that events which straddle window boundaries are still allocated to the window with greatest temporal intersection.

- If no overlap is found, the event is discarded (not assigned to any heartbeat).

After processing all events in this manner, each heartbeat window W_i is associated with a dictionary of sublists, indicating which events of each type fall into that heartbeat. Thus, for heartbeat i, one has:

$$\mathcal{H}_i = \Big\{ \mathrm{Q},\ \mathrm{S1},\ \mathrm{T},\ \mathrm{S2} \Big\}.$$

Association of Physiological Events within a Heartbeat:

A heartbeat ideally consists of one occurrence of each of the following events in a characteristic temporal sequence.

However, due to detection noise, segmentation errors, or other factors, it is possible for a heartbeat window W_i to contain multiple candidates of the same type or to miss a certain event type entirely. The goal of the association step is to select, for each event type, at most one candidate event whose relative timing best satisfies the physiological order. The procedure is as follows:

1. **Candidate Selection**: For each event type $e \in \{\mathrm{Q}, \mathrm{S1}, \mathrm{T}, \mathrm{S2}\}$, if the list $\mathcal{H}_i[e]$ of assigned events is nonempty, select the first candidate in that list.

2. **Temporal Ordering Check**: Once the events are selected, the algorithm checks whether the timing of these events makes sense physiologically. It does this by verifying the following conditions:

   The Q-wave must start before the T-wave ends (to validate the QT interval).

   The Q-wave must start before the center of the S2 sound (used for the QS2 interval).

   The Q-wave must also come before the center of the S1 sound (for the QS1 interval)

   The S1 sound must come before the S2 sound in time (important for the S1S2 interval).

   If all of these conditions are satisfied, the selected events are considered valid and returned. If not, the algorithm gives up on associating these events and returns a result where all the events are marked as absent. This approach ensures that only physiologically meaningful sequences are used, diminishing the impact of false positives returned by the U-Net.

Computation of Cardiac Time Intervals: For each associated combination of events within a heartbeat, compute intervals such as:

- S1–S2 interval (S2 center – S1 center)
- QT interval (T end time – Q start time)
- Q–S2 interval (S2 center – Q start time)
- Q–S1 interval (S1 end time – Q start time)

These reference points were chosen so that intervals more closely aligned with the onset, offset and peak fiducial points.

5.3. Median Heartbeat Extraction Approach

This approach was developed in order to provide a robust solution to segmenting noisy signals by performing segmentation on the median heartbeat template of each recording.

5.3.1 R-peak Detection

Accurate identification of R-peaks in an electrocardiogram (ECG) signal is a critical first step for median heartbeat extraction, since each detected R-peak corresponds to one cardiac cycle. In this work, we attempted to use four existing open source libraries plus a custom Pan–Tompkins implementation, where each library's internal preprocessing routines were used to prepare the signal. Those methods were: custom Pan–Tompkins [41], NeuroKit2 [42], HeartPy [43], BioSPPy [44], WFDB–GQRS [45].

Among these, NeuroKit2 was selected due to its robust performance in R-peak detection and seamless integration within Python-based analysis pipelines. NeuroKit2's algorithms have been validated on benchmark datasets, therefore, to further validate NeuroKit2's performance on the ULSGE ECG signals dataset, a MATLAB-based ECG annotation tool was developed to manually annotate R-peaks. These reference annotations were then used to evaluate the accuracy of the automatic detections. Out of the 84 available ECG recordings, 15 signals were excluded from the evaluation due to poor signal quality and annotation ambiguity. The resulting performance metrics, averaged across the test set, are summarized in Table 5.1.

A tolerance window of 50 milliseconds was used to determine a correct match between manually annotated and automatically detected R-peaks. This threshold is commonly employed in ECG validation studies, as it reflects both the typical precision of human annotation and the physiological variability in QRS onset. It also accounts for minor temporal misalignments due to sampling rate limitations and filtering artifacts. This window

Table 5.1: Overall R-peak detection performance of NeuroKit2 on the ULSGE dataset. A tolerance of 50 ms was used to match detected peaks against manual annotations.

Metric	Value
Total True Positives (TP)	2154
Total False Positives (FP)	118
Total False Negatives (FN)	69
Precision	0.9481
Sensitivity (Recall)	0.9690
F1-Score	0.9584
False Discovery Rate (FDR)	0.0519
Mean Timing Error (s)	0.00204
Standard Deviation of Error (s)	0.00397

ensures a clinically meaningful evaluation of R-peak detection performance without penalizing trivial misalignments.

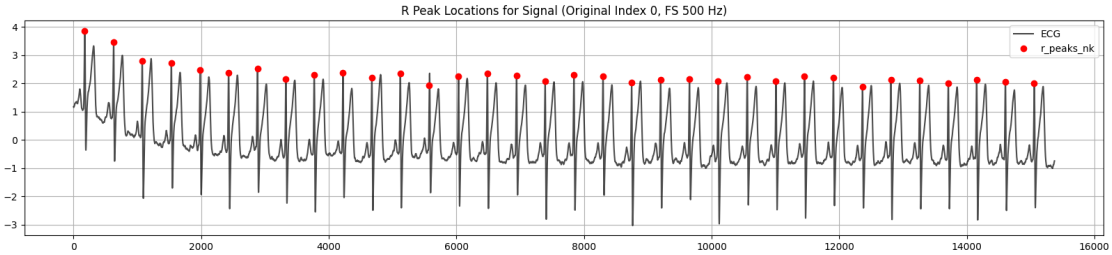

Figure 5.6: Neurokit2's detected R peaks of a ECG recording from the ULSGE dataset

Results and Discussion The results demonstrate strong overall performance, with a precision of 94.81% and recall of 96.90%, leading to an F1-score of 95.84%, which indicates a balanced trade-off between false positives and false negatives. The false discovery rate (FDR) remained low at 5.19%, reflecting the reliability of the detections. Importantly, the mean absolute timing error was approximately 2 ms, with a standard deviation of 4 ms, indicating high temporal accuracy in detecting R-peaks relative to manual annotations.

These metrics justify the use of NeuroKit2 in the preprocessing pipeline, offering a reliable and validated method for R-peak detection that generalizes well to the ULSGE dataset.

5.3.1.1 Median Heartbeat Extraction

Once each ECG and PCG signal has been preprocessed (as described in 5.1) and the R-peaks have been detected, the next step is to align individual cardiac cycles on the R-peak and compute a *median heartbeat*. This subsection describes (1) how individual

ECG beats are extracted and aggregated into a median waveform, and (2) how the PCG median beat is computed using the ECG R-peaks for temporal alignment.

1. ECG Median-Beat Extraction Suppose $\mathcal{R}_i$ is valid for subject i. Let the ECG sampling frequency be F_{ECG} (e.g., $F_{\mathrm{ECG}} = 500\,\mathrm{Hz}$). We define two large time-windows (in seconds), to accommodate the entire hearbeat:

$$T_{\mathrm{pre}} = 0.4\,\mathrm{s}, \qquad T_{\mathrm{post}} = 1.0\,\mathrm{s}.$$

These correspond to a fixed number of samples before and after each R-peak:

$$L_{\mathrm{pre}} = T_{\mathrm{pre}} \times F_{\mathrm{ECG}}, \qquad L_{\mathrm{post}} = T_{\mathrm{post}} \times F_{\mathrm{ECG}}.$$

Thus, for each R-peak $r_{i,k} \in \mathcal{R}_i$, we extract a window of length

$$L_{\mathrm{total}} \;=\; L_{\mathrm{pre}} + L_{\mathrm{post}}$$

samples, centered on $r_{i,k}$.

Stacking these valid beats column-wise yields the *beats matrix* $\mathbf{S}_i$

The *median heartbeat* is computed by taking the sample wise median across all columns of $\mathbf{S}_i$:

$$\mathrm{med}_i[m] = \mathrm{median}\Big\{\, \mathbf{S}_i[m,\,1],\; \mathbf{S}_i[m,\,2],\; \ldots,\; \mathbf{S}_i[m,n] \Big\}, \quad m = 0, 1, \ldots, L_{\mathrm{total}} - 1.$$

Additionally, we calculate the RR intervals (in samples) between successive R-peaks:

$$\Delta r_{i,k} = r_{i,k+1} - r_{i,k}, \quad k = 1, 2, \ldots, K_i - 1,$$

and define the *median RR interval* as

$$\widetilde{RR}_i = \mathrm{median}\Big\{ \Delta r_{i,1},\, \Delta r_{i,2},\, \ldots,\, \Delta r_{i,K_i-1} \Big\}.$$

2. PCG Median-Beat Extraction Using ECG R-Peaks For PCG, we leverage the ECG derived R-peak timings to segment the heart sounds. We fix the same time windows, found for ECG, on the PCG signal and extract the corresponding PCG segments. This is possible because the signals are synchronized.

All valid segments form the PCG *beats matrix*. The median PCG beat is then

$$\text{med}_i^{\text{PCG}}[m] = \text{median}\Big\{ \mathbf{Q}_i[m,\,1],\, \mathbf{Q}_i[m,\,2],\, \ldots,\, \mathbf{Q}_i[m,n] \Big\}, \quad m = 0, \ldots, L_{\text{total}}^{\text{PCG}} - 1.$$

Summary of Median Extraction

- **ECG:** For each subject, extract n windows of length L_{total} samples around each validated R peak, form $\mathbf{S}_i$, and compute $\text{med}_i = \text{median}(\mathbf{S}_i,\ \text{axis} = 1)$. Also compute $\widetilde{RR}_i = \text{median}(\Delta r_{i,k})$.
- **PCG:** Convert ECG R-peak indices to PCG sample indices, extract n windows of length $L_{\text{total}}^{\text{PCG}}$, assemble $\mathbf{Q}_i$, and compute $\text{med}_i^{\text{PCG}} = \text{median}(\mathbf{Q}_i,\ \text{axis} = 1)$.

These median waveforms provide robust, noise resilient templates of the average cardiac cycle in both electrical (ECG) and acoustic (PCG) domains.

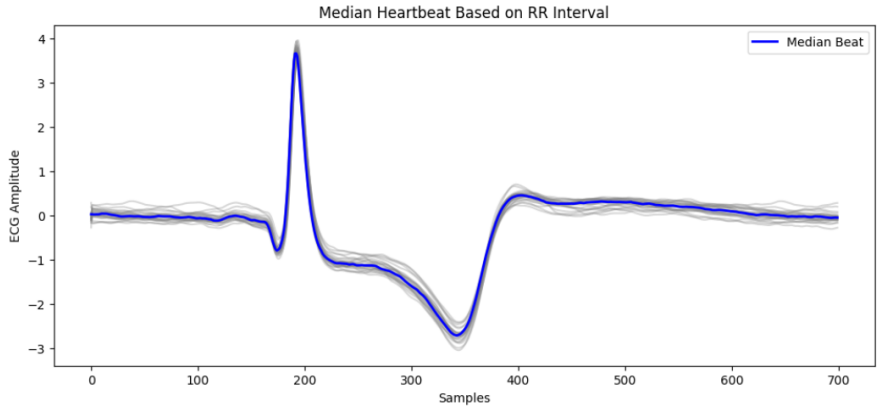

Figure 5.7: Median Heartbeat template of a ECG signal. From the Median Heartbeat approach

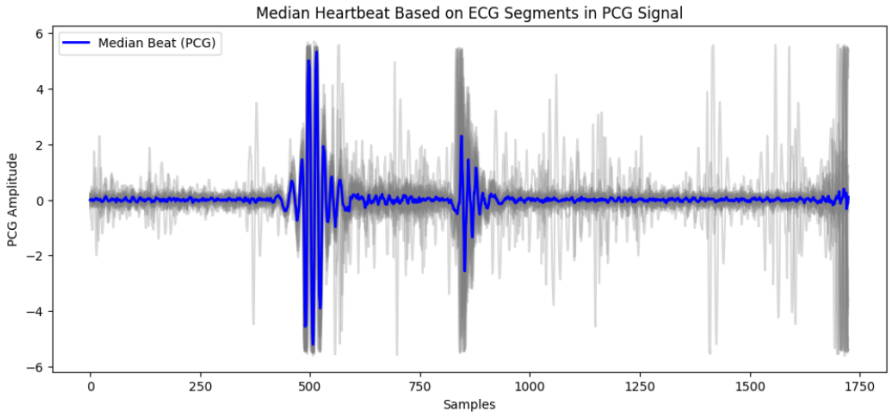

Figure 5.8: Median Heartbeat template of a PCG signal. From the Median Heartbeat approach

5.3.2 Median Heartbeat Segmentation and CTI extraction

Once the template is obtained, the steps that follow are to segment the median heartbeat with the U-Net models and to extract the CTIs from the output probabilities. These steps have already been described in sections 5.2.2 and 5.2.4

6. Results

This chapter's objective is to compare the two segmentation methodologies developed, *Median Heartbeat* and *Beat-by-Beat*, and their ability to accurately estimate the Cardiac Time Intervals.

In section 2.2.4, we discussed the clinical relevance of the intervals LVET (S1S2), QS2, PEP (QS1) and QT as important markers of heart failure. These intervals are directly derived from specific fiducial points identified in synchronized ECG and PCG signals. By detecting these fiducial points, we are able to compute the these intervals.

Since annotations were performed on the median heartbeat template of each recording, the average of each interval estimated by the Beat-by-Beat Approach was calculated in order to compare estimates against annotated values.

Following this, we examine how the derived CTIs relate to patients' ejection fraction (EF) categories, by analysing the statistically significant differences between groups.

6.1. Accuracy of Cardiac Time Intervals) Estimations

The objective of this section is to quantify how well each segmentation approach estimates the CTI's. This is done by comparing predicted versus reference values across multiple statistical metrics. Accuracy is evaluated using:

- Mean Absolute Error (MAE): the average absolute difference between predicted and true interval durations.
- Root Mean Squared Error (RMSE): the square root of the average squared difference, highlighting larger deviations.
- Coefficient of Determination (R^2): indicating the proportion of variance in true values explained by the predictions.
- Pearson and Spearman Correlation Coefficients: measuring linear and rank-based associations, respectively, between predicted and reference intervals.
- Mann–Whitney U Test: used to compare interval distributions between two EF groups (EF2 vs. EF3), to gauge whether the estimation accuracy differs by EF category. The group EF1 was not analysed due to the lack of data points.

Below, we present the results for each interval in tabular form. For clarity, the better-performing method for each metric is highlighted in bold.

6.1.1 S1S2 Interval

Table 6.1: Comparison of S1S2 Interval Analysis Results (all values in milliseconds)

Metric	Median Heartbeat	Beat-by-Beat
Number of recordings	**76**	63
MAE	**16.7**	82.2
RMSE	**23.5**	93.2
R^2	**0.4809**	-4.9551
Pearson correlation (p-value)	**0.7875** ($p < 0.0001$)	0.1552 ($p = 0.2245$)
Spearman correlation (p-value)	**0.7354** ($p < 0.0001$)	0.0069 ($p = 0.9574$)

Scatter Plot Analysis. When comparing the scatter plots of the two estimation methodologies (Median Heartbeat vs. Beat-by-Beat), several key observations can be made:

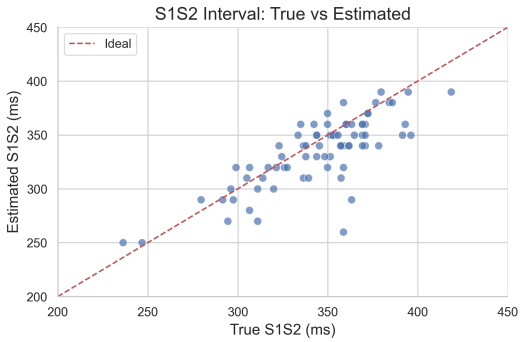

Figure 6.1: Scatter plot of S1S2 intervals extracted with the Median Hearbeat methodology.

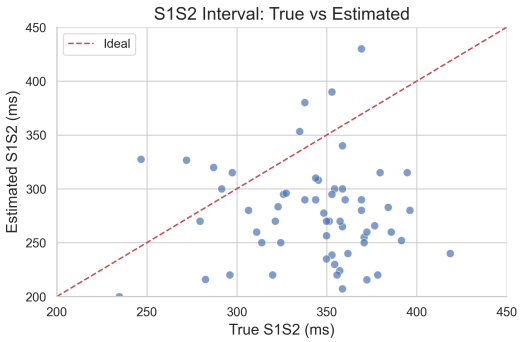

Figure 6.2: Scatter plot of average S1S2 intervals extracted with the Beat-by-Beat methodology.

- In the *Median Heartbeat* approach, the estimated S1S2 intervals cluster closely around the identity line (45°), indicating good agreement with the true values. This method effectively captures the full range of interval durations and can detect this interval in a larger number of recordings (76 compared to 63 for the Beat-by-Beat method).
- In contrast, the *Beat-by-Beat* approach produces estimates that are heavily concentrated around a narrow range (approximately 340–360 ms), failing to capture the variability of the true S1S2 intervals. This results in a clear divergence from the identity line and poorer performance.
- Quantitatively, the **Median Heartbeat** method exhibits a strong and statistically significant correlation with the ground truth, with a Pearson correlation coefficient of

$r = 0.7875$ and Spearman's rank correlation coefficient of $\rho = 0.7354$, both with $p < 0.0001$. The coefficient of determination is $R^2 = 0.4809$, indicating that about 48% of the variance in the true S1S2 values is explained by this method.

- On the other hand, the **Beat-by-Beat** approach shows virtually no linear or monotonic relationship with the true intervals, with Pearson's $r = 0.1552$ ($p = 0.2245$) and Spearman's $\rho = 0.0069$ ($p = 0.9574$). The corresponding $R^2 = -4.9551$ confirms that the estimates are essentially random with respect to the true intervals.

Bland-Altman Plot Analysis. A complementary view is provided by the Bland-Altman analysis:

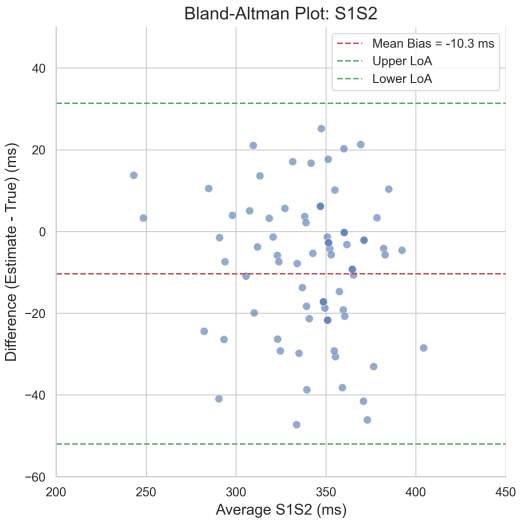

Figure 6.3: Bland-Altman plot of S1S2 intervals extracted with the Median Hearbeat methodology.

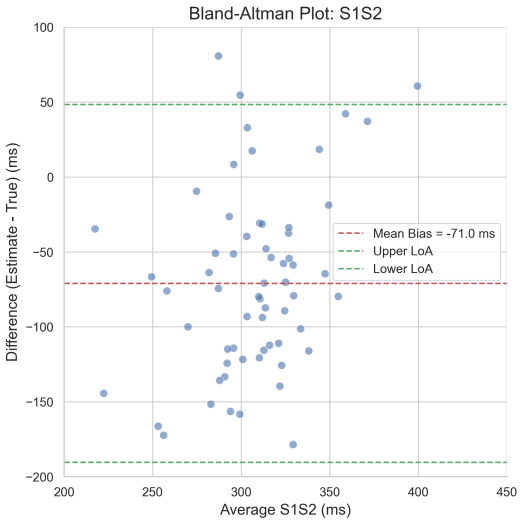

Figure 6.4: Bland-Altman plot of S1S2 intervals extracted with the Beat-by-Beat methodology.

- For the *Median Heartbeat* method, the mean bias is approximately -10.3 ms, indicating that, on average, estimated S1S2 intervals are slightly shorter than the true values. The limits of agreement (LoA) are approximately $[-55$ ms$, +32$ ms$]$. This suggests that most estimates fall within ±50 ms of the true value. There is no obvious trend in the bias across the measurement range, implying that the error is homoscedastic (constant across values).

- In contrast, the *Beat-by-Beat* approach has a much larger bias of -43.1 ms, with limits of agreement spanning approximately $[-400$ ms$, +300$ ms$]$. These wide limits are largely driven by extreme outliers, highlighting poor agreement and high variability in the estimates. This again reinforces the unreliability of the Beat-by-Beat approach when used without the median template alignment.

6.1.2 QT Interval

Table 6.2: Comparison of QT Interval Analysis Results (all values in milliseconds)

Metric	**Median Heartbeat**	**Beat-by-Beat**
Number of recordings	76	**84**
MAE	**27.6**	30.9
RMSE	**39.5**	41.2
R^2	0.1844	**0.2485**
Pearson correlation (p-value)	0.4704 ($p < 0.0001$)	**0.5480** ($p < 0.0001$)
Spearman correlation (p-value)	0.5316 ($p < 0.0001$)	**0.5722** ($p < 0.0001$)

Scatter Plot Analysis. For the QT interval estimation, both the *Median Heartbeat* and *Beat-by-Beat* methods demonstrate some degree of alignment with the ideal identity line. However, a notable amount of dispersion is present in both cases, indicating significant variability in estimation accuracy.

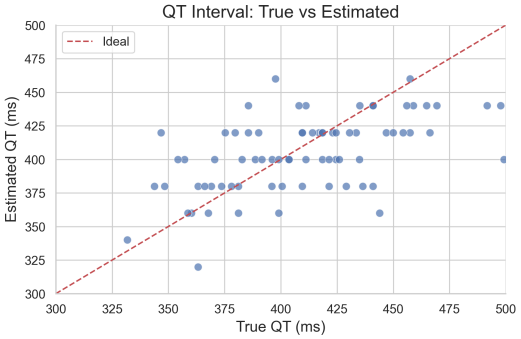

Figure 6.5: Scatter plot of QT intervals extracted with the Median Hearbeat methodology.

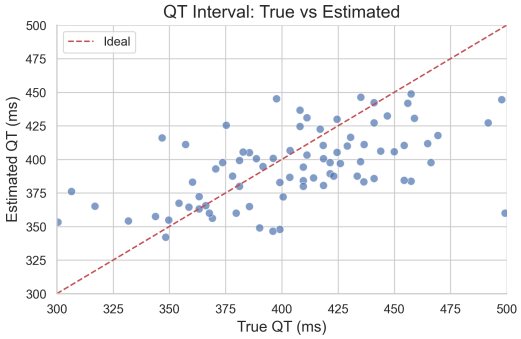

Figure 6.6: Scatter plot of average QT intervals extracted with the Beat-by-Beat methodology.

- The *Median Heartbeat* approach yields a Pearson correlation of $r = 0.4704$ and Spearman's $\rho = 0.5316$, both with $p < 0.0001$, signifying a moderate and statistically significant linear and monotonic relationship between the estimated and true QT intervals.
- The *Beat-by-Beat* method shows slightly stronger performance, with Pearson's $r = 0.5480$ and Spearman's $\rho = 0.5722$, again with $p < 0.0001$. The coefficient of determination $R^2 = 0.2485$ suggests that nearly 25% of the variance in the true QT intervals is explained by the Beat-by-Beat estimates, compared to approximately 18% ($R^2 = 0.1844$) for the Median Heartbeat approach.
- In both methods, the scatter plots indicate a trend following the identity line, but with relatively high residual variability. This suggests that while both models capture the general relationship, the precision of individual estimates remains limited.

- Since the QT interval is derived from ECG signals, which generally have higher signal-to-noise ratio and less variability than PCG, the Beat-by-Beat method is better able to extract QT intervals from more recordings (84 vs. 76). This reflects the advantage of Beat-by-Beat estimation in processing ECG-derived intervals.

Bland-Altman Plot Analysis. The Bland-Altman analysis further illustrates the estimation characteristics:

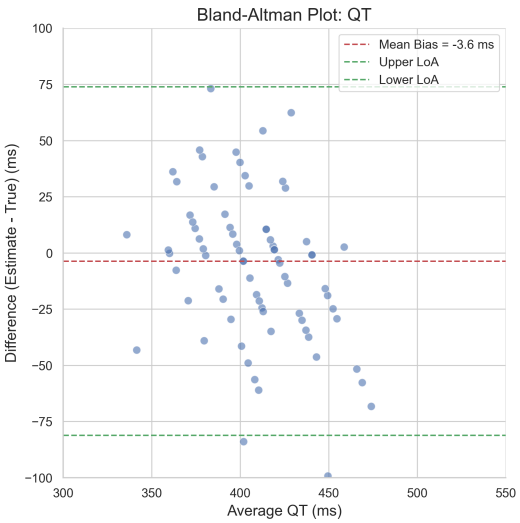

Figure 6.7: Bland-Altman plot of QT intervals extracted with the Median Heartbeat methodology.

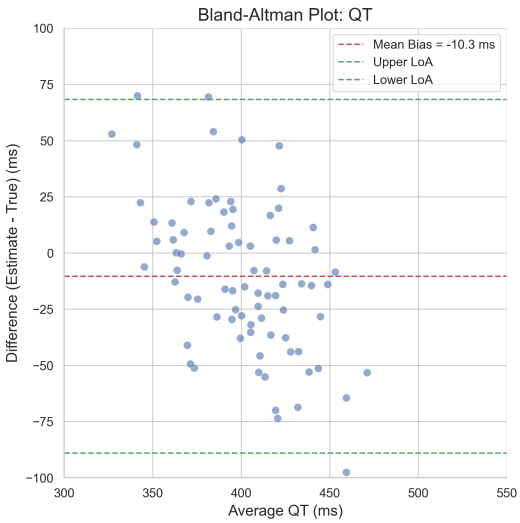

Figure 6.8: Bland-Altman plot of average QT intervals extracted with the Beat-by-Beat methodology.

- For the *Median Heartbeat* method, the mean bias is approximately -3.6 ms, indicating that the QT interval is, on average, slightly underestimated. The limits of agreement (LoA) are approximately $[-75$ ms, $+75$ ms$]$, reflecting a symmetric and relatively narrow distribution of errors around the mean.

- The *Beat-by-Beat* method exhibits a slightly larger mean bias of -10.3 ms, again indicating a small tendency to underestimate the true QT interval. The limits of agreement are similar to those of the Median Heartbeat method, also roughly in the range of $[-75$ ms, $+75$ ms$]$. This implies comparable spread in the estimation errors across both methodologies.

- Neither method shows a clear pattern of increasing or decreasing error across the range of QT values, suggesting that the bias is roughly consistent across the measurement spectrum (i.e., homoscedasticity).

6.1.3 QS2 Interval

Table 6.3: Comparison of QS2 Interval Analysis Results (all values in milliseconds)

Metric	Median Heartbeat	Beat-by-Beat
Number of recordings	75	75
MAE	**42.0**	178.4
RMSE	**44.8**	216.1
R^2	-1.2279	**-33.9816**
Pearson correlation (p-value)	**0.8091** ($p < 0.0001$)	-0.0748 ($p = 0.5235$)
Spearman correlation (p-value)	**0.7705** ($p < 0.0001$)	0.1151 ($p = 0.3256$)

Scatter Plot Analysis. In the estimation of the QS2 interval, both methods reveal limited agreement with the identity line, though their behaviors differ:

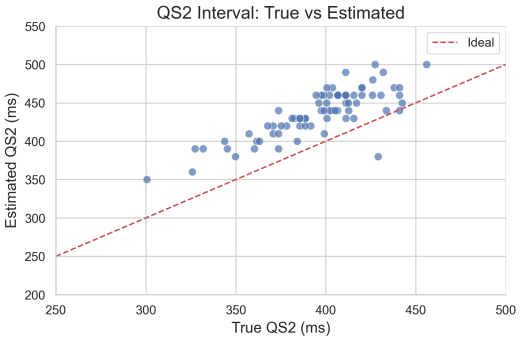

Figure 6.9: Scatter plot of QS2 intervals extracted with the Median Heartbeat methodology.

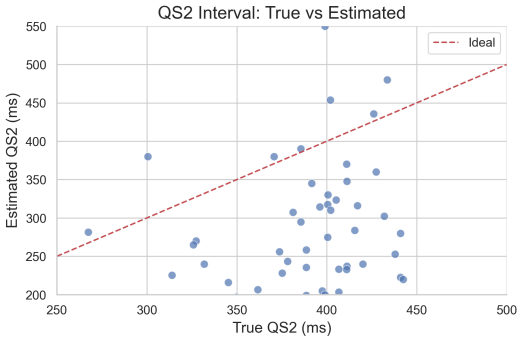

Figure 6.10: Scatter plot of average QS2 intervals extracted with the Beat-by-Beat methodology.

- The *Median Heartbeat* approach exhibits a strong linear trend with moderate dispersion. The Pearson correlation coefficient is $r = 0.8091$ with $p < 0.0001$, and Spearman's rank correlation is $\rho = 0.7705$ with $p < 0.0001$, indicating a statistically significant and strong linear and monotonic relationship with the true QS2 intervals.
- In contrast, the *Beat-by-Beat* method shows no meaningful correlation, with Pearson's $r = -0.0748$ ($p = 0.5235$) and Spearman's $\rho = 0.1151$ ($p = 0.3256$), suggesting the estimates are essentially uncorrelated with the true values. The scatter plot does not follow the identity line and lacks clear structure.
- The coefficient of determination R^2 indicates poor fits in both cases: $R^2 = -1.2279$ for Median Heartbeat and an even worse $R^2 = -33.9816$ for Beat-by-Beat, highlighting that these models perform worse than simply predicting the mean QS2 interval.

Bland-Altman Plot Analysis. The Bland-Altman plots highlight the systematic bias and error magnitude of each method:

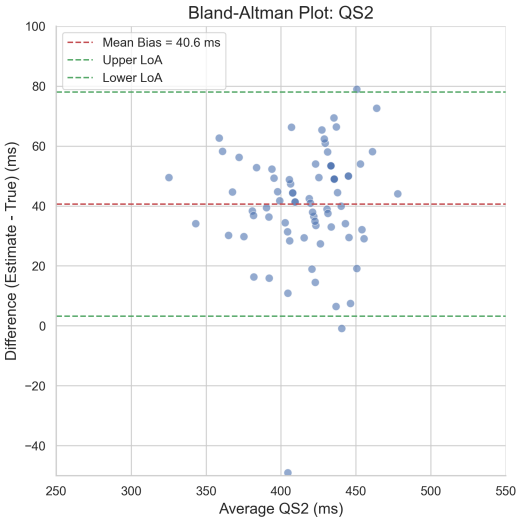

Figure 6.11: Bland-Altman plot of QS2 intervals extracted with the Median Heartbeat methodology.

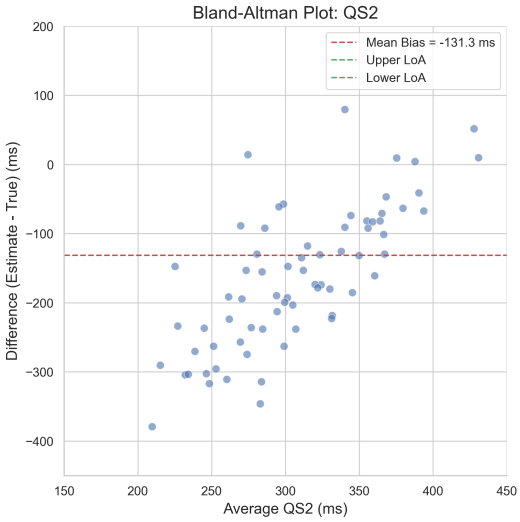

Figure 6.12: Bland-Altman plot of average QS2 intervals extracted with the Beat-by-Beat methodology.

- For the *Median Heartbeat*, the mean bias is approximately $+40.6$ ms, suggesting a consistent overestimation of the QS2 interval. The limits of agreement range from approximately -40 ms to $+120$ ms, indicating moderate variability, with most estimates differing from the ground truth by less than 120 ms in either direction.

- The *Beat-By-Beat* approach shows a considerably larger mean bias of $+131.3$ ms, again in the direction of overestimation. The limits of agreement are extremely wide, ranging from approximately -600 ms to $+600$ ms. This massive spread is indicative of highly inconsistent estimates and the influence of outliers.

- The variability in both methods does not exhibit a strong funnel shape, though the large spread in the Beat-by-Bea approach suggests unreliable performance across the full range of QS2 intervals.

6.1.4 QS1 Interval

Table 6.4: Comparison of QS1 Interval Analysis Results (all values in milliseconds)

Metric	Median Heartbeat	Beat-by-Beat
Number of recordings	82	83
MAE	**53.6**	189.4
RMSE	**62.8**	222.8
R^2	**-9.8101**	-132.1682
Pearson correlation (p-value)	0.1037 ($p = 0.3537$)	0.1531 ($p = 0.1671$)
Spearman correlation (p-value)	**0.4014** ($p = 0.0002$)	0.1013 ($p = 0.3622$)

Scatter Plot Analysis. In the estimation of the QS1 interval, both methods over estimate QS1 and are poorly aligned with the identity line:

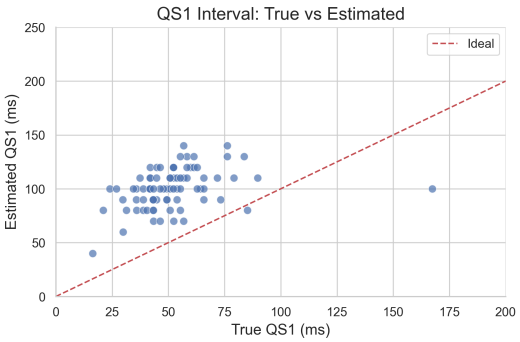

Figure 6.13: Scatter plot of QS1 intervals extracted with the Median Heartbeat methodology.

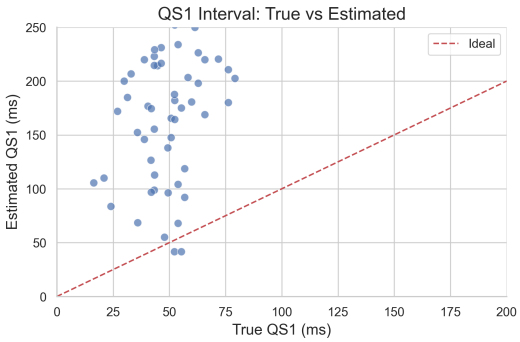

Figure 6.14: Scatter plot of average QS1 intervals extracted with the Beat-by-Beat methodology.

- The *Median Heartbeat* approach does not exhibit a strong linear trend and shows notable dispersion. Most points lie above the 45° line, indicating a consistent overestimation of the QS1 interval. The Pearson correlation is $r = 0.1037$ with $p = 0.3537$, and Spearman's $\rho = 0.4014$ with $p = 0.0002$, indicating a weak linear but moderate and statistically significant monotonic relationship with the reference values.

- The *Beat-by-Beat* method also shows no meaningful correlation, with Pearson's $r = 0.1531$ ($p = 0.1671$) and Spearman's $\rho = 0.1013$ ($p = 0.3622$), suggesting the estimates are essentially uncorrelated with the true values. The scatter plot does not follow the identity line and lacks a clear linear or monotonic structure.

- The coefficient of determination R^2 reinforces this poor fit: $R^2 = -9.8101$ for the Median Heartbeat approach, and $R^2 = -132.1682$ for the Beat-by-Beat method, indicating that both models perform substantially worse than simply predicting the mean QS1 interval.

Bland-Altman Plot Analysis. The Bland-Altman plots highlight the systematic bias and error magnitude of each method:

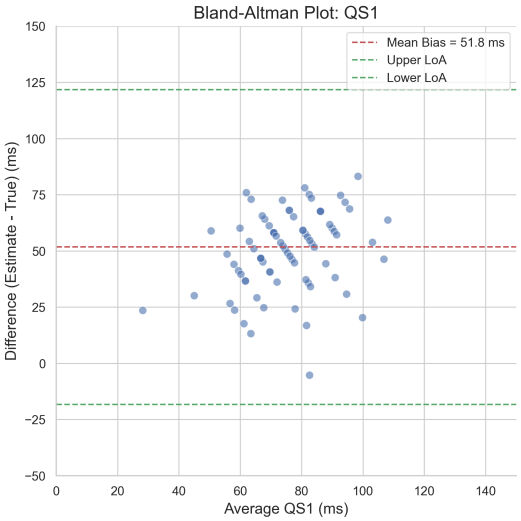

Figure 6.15: Bland-Altman plot of QS1 intervals extracted with the Median Heartbeat methodology.

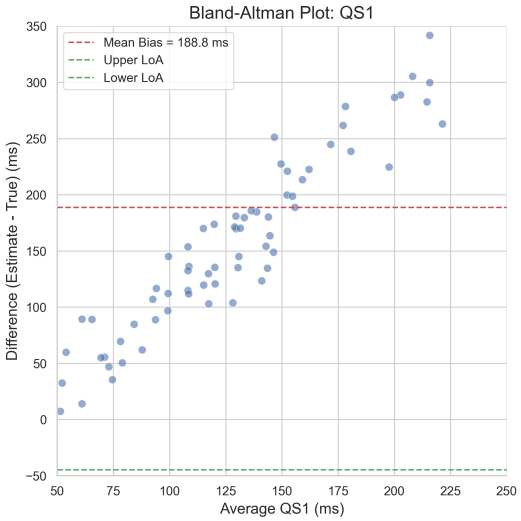

Figure 6.16: Bland-Altman plot of average QS1 intervals extracted with the Beat-by-Beat methodology.

- For the *Median Heartbeat* approach, the mean bias is approximately $+51.8$ ms, indicating a consistent overestimation of the QS1 interval. The limits of agreement span from roughly -25 ms to $+150$ ms, reflecting moderate variability in the estimates.
- The *Beat-by-Beat* method exhibits a substantially larger mean bias of $+188.8$ ms, also indicating systematic overestimation. The limits of agreement are extremely wide, ranging from approximately 0 ms to 430.9 ms, suggesting very high variability. This method consistently overestimates the QS1 interval across recordings.

6.2. Ejection Fraction - Based Analysis

Next, we investigate whether the estimated CTIs differ significantly across EF categories. Patients were stratified into three EF groups:

- **EF1:** EF < 40%
- **EF2:** 40% ≤ EF ≤ 50%
- **EF3:** EF > 50%

For each interval (S1S2, QS1, QS2, QT), we compared distributions between EF2 and EF3 using the Mann–Whitney U test. This nonparametric test was chosen due to non-normality in several interval distributions and unequal sample sizes between groups.

The test was conducted to compare interval estimates between EF group 2 (borderline, EF 40–50%) and EF group 3 (preserved EF, EF $>$ 50%). Group 1 (reduced EF, EF $<$ 40%) was excluded from the statistical comparison due to insufficient sample size ($n = 4$), which would compromise the reliability of nonparametric tests.

6.2.1 S1S2 Interval

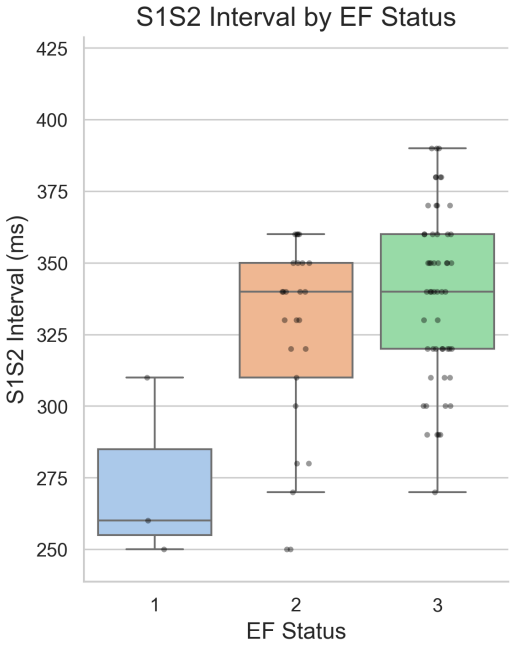

Figure 6.17: Boxplot of estimated S1S2 intervals, with the Median Heartbeat methodology, per Ejection Fraction group

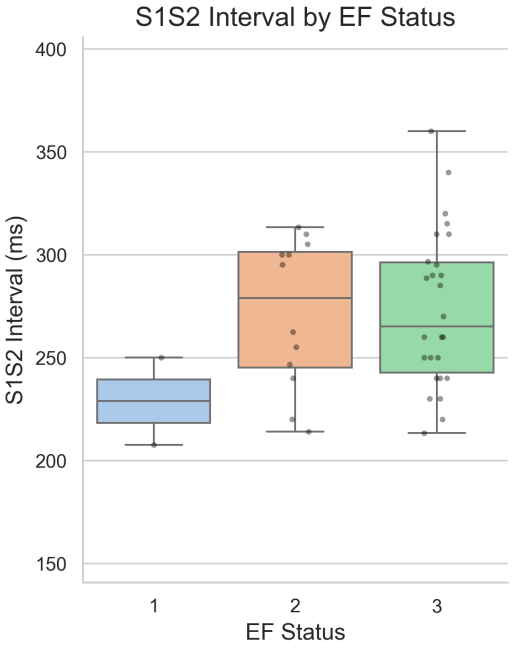

Figure 6.18: Boxplot of estimated S1S2 intervals, with the Beat-by-Beat methodology, per Ejection Fraction group

Table 6.5: Mann-Whitney Test: S1S2 Interval

Metric	Median Heartbeat	Beat-by-Beat
Mann-Whitney U (EF2 vs EF3)	**488.0** ($p = 0.2490$)	374.0 ($p = 0.4485$)

For the S1S2 interval, the *Median Heartbeat* method yielded a Mann–Whitney U statistic of $U = 488.0$ with $p = 0.2490$, suggesting a trend toward differentiation between EF groups 2 and 3, though not reaching statistical significance at the conventional $\alpha = 0.05$ threshold. In contrast, the *Beat-by-Beat* approach failed to show any meaningful separation between the groups ($U = 374.0$, $p = 0.4485$), consistent with

its poor correlation metrics and large error distribution for this interval. These findings reinforce that the Median Heartbeat method retains some sensitivity to physiological differences in EF status, even when not strictly significant, whereas the Beat-by-Be's interval estimates are effectively random with respect to EF grouping.

6.2.2 QT Interval

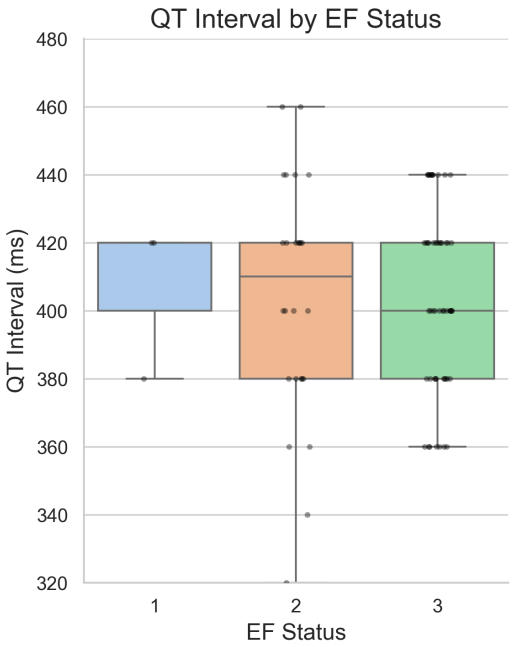

Figure 6.19: Boxplot of estimated QT intervals, with the Median Heartbeat methodology, per Ejection Fraction group.

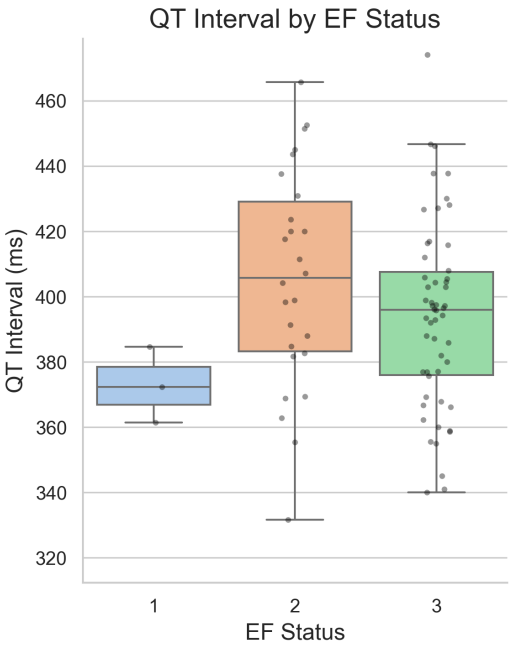

Figure 6.20: Boxplot of estimated QT intervals, with the Beat-by-Beat methodology, per Ejection Fraction group

Table 6.6: Mann-Whitney Test: QT Interval

Metric	**Median Heartbeat**	**Beat-by-Beat**
Mann-Whitney U (EF2 vs EF3)	609.0 ($p = 0.8056$)	**864.5** ($p = 0.0961$)

For QT interval estimation, neither method achieved statistical significance, though the Beat-by-Beat approach came closer ($p = 0.0961$) compared to the Median Heartbeat method ($p = 0.8056$). These results reinforce the robustness of the Median Heartbeat method in identifying subtle clinical differences in cardiac function, particularly for QS2.

6.2.3 QS2 Interval

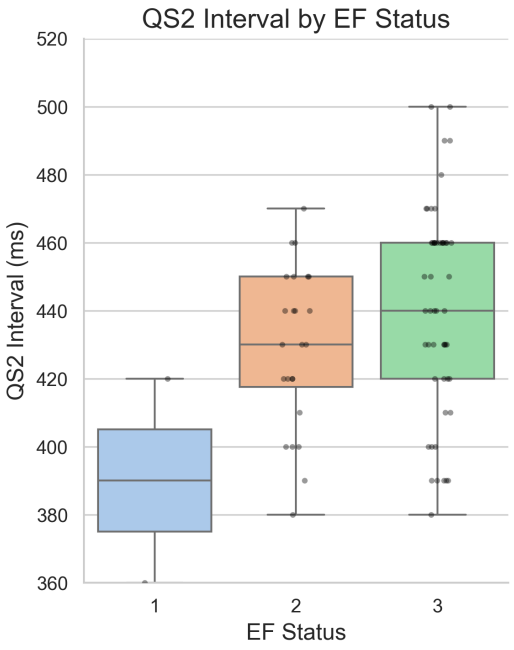

Figure 6.21: Boxplot of estimated QS2 intervals, with the Median Heartbeat methodology, per Ejection Fraction group.

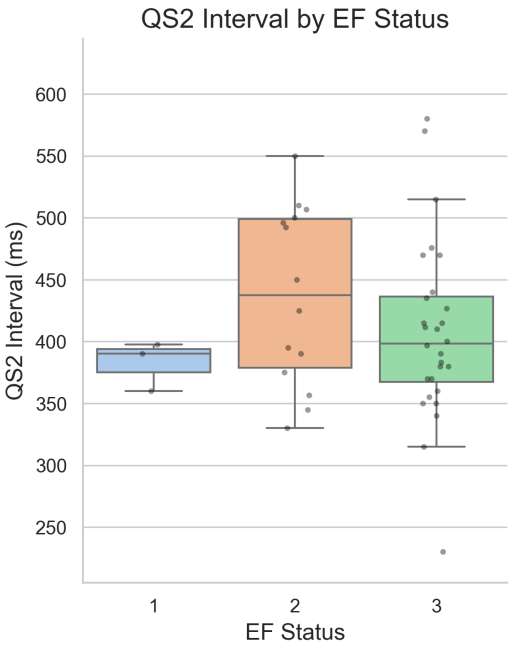

Figure 6.22: Boxplot of estimated QS2 intervals, with the Beat-by-Beat methodology, per Ejection Fraction group

Table 6.7: Comparison of QS2 Interval Analysis Results (all values in milliseconds)

Metric	**Median Heartbeat**	**Beat-by-Beat**
Mann-Whitney U (EF2 vs EF3)	**437.5** ($p = 0.1362$)	529.5 ($p = 0.7869$)

The *Median Heartbeat* approach almost yielded a statistically significant difference in QS2 intervals between groups 2 and 3 ($U = 437.5$, $p = 0.1362$). In contrast, the *Beat-by-Beat* method did not show a significant group separation ($U = 529.5$, $p = 0.7869$), reflecting its poorer correlation with EF-based physiological differences.

6.2.4 QS1 Interval

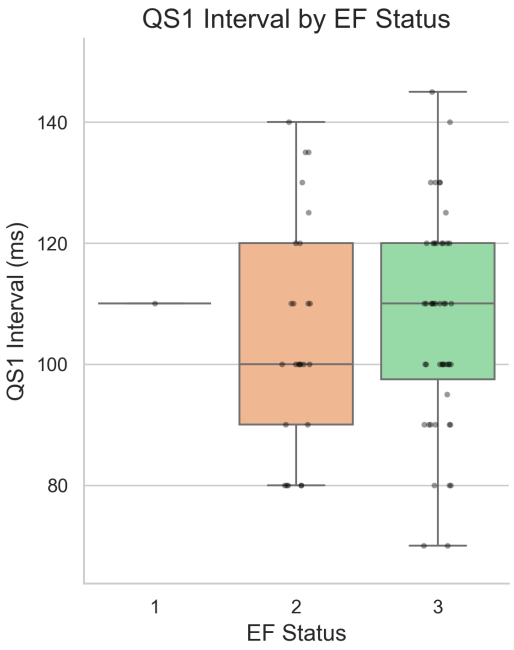

Figure 6.23: Boxplot of estimated QS1 intervals, with the Median Heartbeat methodology, per Ejection Fraction group.

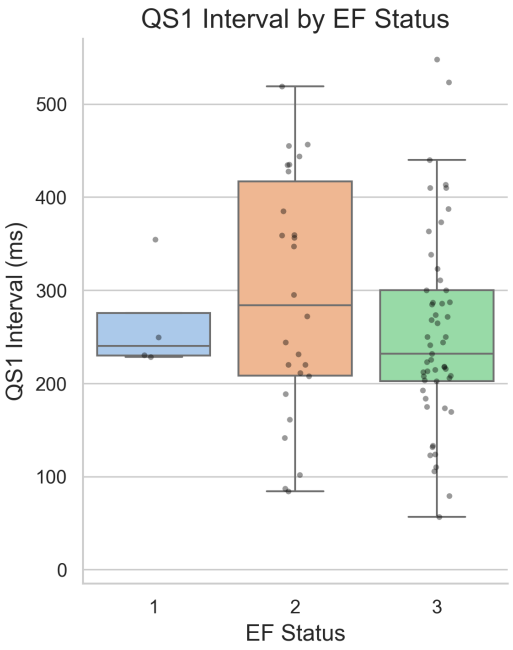

Figure 6.24: Boxplot of estimated QS1 intervals, with the Beat-by-Beat methodology, per Ejection Fraction group

Table 6.8: Comparison of QS1 Interval Analysis Results (all values in milliseconds)

Metric	**Median Heartbeat**	**Beat-by-Beat**
Mann-Whitney U (EF2 vs EF3)	571.500 ($p = 0.2740$)	**821.0000** ($p = 0.1701$)

The *Median Heartbeat* approach did not yield a statistically significant difference in QS1 intervals between EF groups 2 and 3 ($U = 571.5$, $p = 0.2740$), suggesting limited sensitivity to EF-related distinctions. Similarly, the *Beat-by-Bea* method also failed to show significant group separation ($U = 821.0$, $p = 0.1701$), indicating that neither method reliably captures QS1 variations across EF categories in this comparison.

7. Discussion

This work investigated two distinct methodologies for estimating CTIs from non-invasive signals: the *Median Heartbeat* approach and a deep learning-based *Beat-by-Beat* model. Both methods aimed to extract clinically relevant intervals such as S1S2, QT, QS1 and QS2 from ECG and PCG signals. Each method has its own set of strengths and limitations, particularly when applied to multimodal physiological data.

7.1. Challenges in CTI Estimation from ECG and PCG

One of the main challenges encountered lies in the intrinsic properties of the signals involved. ECG and PCG signals were processed independently, and many CTIs—such as QS2—span across both modalities. As a result, any degradation in quality in either the ECG or PCG segments can propagate error to the final interval estimate.

The PCG signal, in particular, presents greater difficulty. Acoustic signals are more susceptible to environmental noise, microphone placement, and anatomical variability. Moreover, heart sounds such as S1 and S2 have less clearly defined morphology and onset/offset characteristics compared to ECG events like QRS complexes. This ambiguity increases both manual labeling variability and model prediction uncertainty.

7.2. Performance Comparison and Interpretation

The Median Heartbeat method significantly outperformed the Beat-by-Beat approach in intervals derived from PCG signals, such as S1S2 and QS2. PCG signals are inherently more variable and susceptible to noise compared to ECG, which challenges beat-level interval estimation. By aggregating information over multiple beats, the Median Heartbeat method provides more robust and consistent estimates, effectively mitigating the impact of non-stationary noise and morphological variability typical in PCG data.

Specifically, QS2 is a hybrid interval derived from both PCG and ECG signals. Due to the complex and noisy nature of PCG and its interaction with ECG, the Beat-by-Beat method struggled with noise and variability, resulting in poorer performance as reflected by large errors and weak correlations. In contrast, the Median Heartbeat

method yielded markedly lower error metrics and stronger correlations, demonstrating its advantage in handling multi-modal signals through averaging.

Conversely, the Beat-by-Beat method performed better in ECG-based intervals such as QT, which benefit from the higher signal-to-noise ratio and more consistent morphology of ECG signals. This is exemplified by the Beat-by-Beat method extracting QT intervals from a greater number of recordings (84 vs. 76 for the Median Heartbeat) and showing stronger statistical fits (higher R^2 and correlation values). This suggests that beat-level analysis can leverage the stability of ECG waveforms more effectively than in PCG-derived intervals.

When comparing QS1 and QS2 intervals, larger errors were observed for QS1 across both methods. This can be explained by the acoustic properties of the underlying heart sounds: S1 sounds tend to be more variable, have less distinct features, and are more vulnerable to noise and interference than S2 sounds. Such variability particularly affects beat-by-beat estimates, where noise accumulates and causes greater inaccuracy. The Median Heartbeat method partially offsets this challenge by averaging over multiple beats, resulting in relatively better—though still limited—performance for QS1.

In summary, the results highlight the critical importance of signal characteristics in interval estimation. The Median Heartbeat method's strength lies in its robustness to noisy, non-stationary PCG signals and hybrid intervals, while the Beat-by-Beat method capitalizes on the cleaner, more structured ECG signals to produce more accurate beat-level interval estimates.

7.3. EF Group Stratification

Another point of evaluation was the ability of both methods to capture clinically relevant stratifications, particularly between EF groups. The Median Heartbeat method revealed significant group separation for the QS2 interval and showed a trend in the right direction for S1S2, aligning with known pathophysiological patterns (e.g., prolonged QS2 in reduced EF). The Beat-by-Beat approach, however, failed to distinguish EF groups across all intervals, highlighting its lack of physiological sensitivity in its current form.

These findings are further contextualized by results obtained from manual annotations [36]. This analysis demonstrated a significant difference in the QS2 interval

across the three EF groups (Kruskal-Wallis p = 0.008), as well as for the S1S2 interval (p = 0.009), reinforcing the relevance of these markers for stratification. However, pairwise comparisons between the mid-range (EF 40–49%) and preserved EF (≥50%) groups did not reach significance (QS2 (p = 0.159) and S1S2 (p = 0.491)) suggesting that these intervals may be more sensitive in distinguishing reduced EF rather than more subtle gradations.

7.4. Limitations and Future Directions

Several limitations emerged during this work. First, the time resolution of 20 ms used in the Beat-by-Beat model restricted its ability to estimate short-duration intervals with sufficient precision. This limitation is especially critical for the smallest interval studied, QS1, where clinical significance depends on capturing subtle temporal variations. Training with higher temporal resolution, though computationally more demanding, could therefore improve accuracy and clinical relevance. Second, the independent processing of ECG and PCG signals limited the potential synergy across modalities; a multimodal network employing early or mid-level fusion strategies might better exploit complementary information and improve overall performance.

Additionally, dataset limitations—such as the small sample size in the reduced EF group (EF1, $n = 4$)—prevented a full clinical stratification. Future work should aim to collect more balanced and higher-quality datasets to enable more robust statistical comparisons and potentially explore transfer learning or domain adaptation techniques to generalize better across diverse patient groups.

Finally, it is important to highlight that the Beat-by-Beat approach enables the calculation of instantaneous intervals, unlike the Median Heartbeat approach. This capability is particularly valuable for computing corrected intervals, which are more clinically relevant and show stronger correlations with heart failure indicators such as ejection fraction (EF). Enhancing the Beat-by-Beat model could therefore support more insightful clinical analyses and improve the interpretability of CTI-based biomarkers.

7.5. Conclusion

In this work, we developed and evaluated two approaches for the automatic extraction of CTIs: the Median Heartbeat method and the Beat-by-Beat approach. The Median Heartbeat approach was designed to provide a robust estimation of CTIs

on a representative cardiac cycle, while the Beat-by-Beat approach offers a simpler, heuristic-free alternative based on a per-cycle basis. From both methods, we obtained accuracy and LVEF correlation measures, to acess their revelance.

In conclusion, the Median Heartbeat approach remains a reliable and interpretable baseline for CTI estimation, especially for PCG-derived intervals. The Beat-by-Beat model shows promise for ECG-based interval estimation but requires methodological refinements to handle multimodal signals effectively. Future work may benefit from combining the strengths of both approaches—leveraging model-based learning for robust ECG segmentation and template-based refinement for PCG events—to improve performance across the full spectrum of cardiac interval estimation.

References

[1] World Health Organization, “Cardiovascular diseases (CVDs),” https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), 2019, accessed: 2025-06-04. [Cited on page 1.]

[2] G. Savarese and L. H. Lund, “Global public health burden of heart failure: An updated review,” *Cardiac Failure Review*, vol. 3, no. 1, pp. 7–11, 2017. [Cited on page 1.]

[3] R. P. Paiva, P. Carvalho, R. Couceiro, J. Henriques, M. Antunes, I. Quintal, and J. Muehlsteff, “Beat-to-beat systolic time-interval measurement from heart sounds and ECG,” *Physiol. Meas.*, vol. 33, no. 2, pp. 177–194, Feb. 2012. [Cited on pages iv, vi, 1, 7, 17, and 18.]

[4] M. Wu, Z. Yao, J. Zhao, E. Feng, and T. Wang, “Quality of echocardiograms is operator dependent and can vary across patients and equipment,” *Frontiers in Cardiovascular Medicine*, vol. 8, p. 648877, 2021. [Cited on page 1.]

[5] Z. He, X. Zhang, and L. Chen, “Ecg interpretation: Clinical relevance, challenges, and advances,” *Cardiology Discovery*, vol. 2, no. 4, p. 39, 2020. [Cited on page 1.]

[6] B. Zhu, Z. Zhou, S. Yu, X. Liang, Y. Xie, and Q. Sun, “Review of phonocardiogram signal analysis: Insights from the physionet/cinc challenge 2016 database,” *Electronics*, vol. 13, no. 16, p. 3222, 2024. [Cited on page 1.]

[7] M. Klum, M. Urban, T. Tigges, A.-G. Pielmus, A. Feldheiser, T. Schmitt, and R. Orglmeister, “Wearable cardiorespiratory monitoring employing a multimodal digital patch stethoscope: Estimation of ecg, pep, lvet and respiration using a 55 mm single-lead ecg and phonocardiogram,” *Sensors*, vol. 20, no. 7, p. 2033, 2020. [Cited on pages iv, vi, 1, 17, and 20.]

[8] T. A. McDonagh and et al., “2023 focused update of the 2021 esc guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the european society of cardiology (esc) with the special contribution of the heart failure association (hfa) of the esc,” *European*

Heart Journal, vol. 44, no. 37, pp. 3627–3639, 08 2023. [Online]. Available: https://doi.org/10.1093/eurheartj/ehad195 [Cited on pages vi and 4.]

[9] P. A. Heidenreich and et al., "2022 aha/acc/hfsa guideline for the management of heart failure: A report of the american college of cardiology/american heart association joint committee on clinical practice guidelines," *Circulation*, vol. 145, no. 18, pp. e895–e1032, 2022. [Online]. Available: https://www.ahajournals.org/doi/abs/10.1161/CIR.0000000000001063 [Cited on page 4.]

[10] A. Kosaraju, A. Goyal, Y. Grigorova, and A. N. Makaryus, *Left Ventricular Ejection Fraction*, april 24, 2023 ed. Treasure Island (FL): StatPearls Publishing, 2023, available from: https://www.ncbi.nlm.nih.gov/books/NBK459131/. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK459131/ [Cited on pages 4 and 5.]

[11] A. S. Alhakak, F. J. Olsen, K. G. Skaarup, T. Biering-Sørensen, and R. Møgelvang, "Changes in cardiac time intervals over a decade and the risk of incident heart failure: The copenhagen city heart study," *International Journal of Cardiology*, vol. 386, pp. 141–148, 2023. [Cited on page 5.]

[12] S. Hassan and P. Turner, "Systolic time intervals: a review of the method in the non-invasive investigation of cardiac function in health, disease and clinical pharmacology," *Postgrad. Med. J.*, vol. 59, no. 693, pp. 423–434, Jul. 1983. [Cited on page 5.]

[13] R. P. Lewis, S. E. Rittogers, W. F. Froester, and H. Boudoulas, "A critical review of the systolic time intervals," *Circulation*, vol. 56, no. 2, pp. 146–158, Aug. 1977. [Cited on page 5.]

[14] P. Dehkordi, F. Khosrow-Khavar, M. Di Rienzo, O. T. Inan, S. E. Schmidt, A. P. Blaber, K. Sørensen, J. J. Struijk, V. Zakeri, P. Lombardi, M. M. H. Shandhi, M. Borairi, J. M. Zanetti, and K. Tavakolian, "Comparison of different methods for estimating cardiac timings: A comprehensive multimodal echocardiography investigation," *Frontiers in Physiology*, vol. 10, p. 1057, 2019. [Online]. Available: https://doi.org/10.3389/fphys.2019.01057 [Cited on pages 6, 10, and 17.]

[15] M. Michail and A. J. Brown, "Physiology of the normal heart," *Medicine (Abingdon)*, vol. 46, no. 8, pp. 427–430, Aug. 2018. [Cited on pages vi and 6.]

[16] X. Y. Author *et al.*, “Review of phonocardiogram signal analysis: Insights from the physionet/cinc challenge 2016 database,” *Electronics*, 2023. [Cited on page 7.]

[17] S. Kumar *et al.*, “Localization and classification of heart beats in phonocardiography signals—a comprehensive review,” *EURASIP Journal on Advances in Signal Processing*, 2018, reports S1 peak around 50 Hz and most content ≤150 Hz; recommends ≥1 kHz sampling. [Cited on page 7.]

[18] Tang *et al.*, “Pcg classification using multidomain features and svm classifier,” *BioMed Research International*, 2018, frequency features up to 790 Hz extracted; typical PCG main band ≤600 Hz. [Cited on page 7.]

[19] H. T. et al., “Classification of heart sounds using convolutional neural network,” *Applied Sciences*, 2019, used sampling at 2000 Hz dataset. [Cited on page 7.]

[20] L. F. Ekenberg, D. E. Høfsten, S. M. Rasmussen *et al.*, “Wireless single lead versus standard 12 lead ecg, for st segment deviation during adenosine cardiac stress scintigraphy,” *Sensors (Basel)*, vol. 23, no. 6, p. 2962, 2023. [Cited on page 8.]

[21] T. Health, “12 lead ecg vs. single lead ecg: The superior choice,” *Tricog Blog*, 2023, reports sensitivity 44% (single lead) vs. 79% (12 lead) for ST elevation. [Cited on page 8.]

[22] A. Mohamoud, J. Jensen, and K. G. Buda, “Consumer grade wearable cardiac monitors: What they do well, and what needs work,” *Cleveland Clinic Journal of Medicine*, vol. 91, no. 1, p. 23–29, 2024. [Cited on page 8.]

[23] E. Vergara-Escudero, S. Vieau, and P. A. Iaizzo, “Basic ECG theory, 12-lead recordings and their interpretation,” in *Handbook of Cardiac Anatomy, Physiology, and Devices*. Cham: Springer Nature Switzerland, 2024, pp. 317–332. [Cited on page 8.]

[24] K. Tavakolian, “Systolic time intervals and new measurement methods,” *Cardiovasc. Eng. Technol.*, vol. 7, no. 2, pp. 118–125, Jun. 2016. [Cited on page 10.]

[25] M. Roos, S. Toggweiler, M. Zuber, P. Jamshidi, and P. Erne, “Acoustic cardiographic parameters and their relationship to invasive hemodynamic

measurements in patients with left ventricular systolic dysfunction," *Congestive Heart Failure*, vol. 12, no. s1, pp. 19–24, 2006. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1527-5299.2006.05769.x [Cited on pages vi and 11.]

[26] S. Wang, Y.-Y. Lam, M. Liu, F. Fang, J. Wang, Q. Shang, J.-P. Sun, J. E. Sanderson, and C.-M. Yu, "Acoustic cardiography helps to identify heart failure and its phenotypes," *International Journal of Cardiology*, vol. 167, no. 3, pp. 681–686, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167527312002720 [Cited on page 12.]

[27] M. Zuber, P. Kipfer, and C. A. Jost, "Systolic dysfunction: Correlation of acoustic cardiography with doppler echocardiography," *Congestive Heart Failure*, vol. 12, no. s1, pp. 14–18, 2006. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0889-7204.2006.05768.x [Cited on page 12.]

[28] N. Giordano, S. Rosati, G. Balestra, and M. Knaflitz, "A wearable multi-sensor array enables the recording of heart sounds in homecare," *Sensors*, vol. 23, no. 13, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/13/6241 [Cited on page 12.]

[29] H. Boudoulas, Y. H. Sohn, W. O'neill, R. Brown, and A. M. Weissler, "The qt > qs2 syndrome: A new mortality risk indicator in coronary artery disease," *The American Journal of Cardiology*, vol. 50, no. 6, pp. 1229–1235, 1982. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0002914982904544 [Cited on page 12.]

[30] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The prisma 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, 2021. [Online]. Available: https://www.bmj.com/content/372/bmj.n71 [Cited on page 13.]

[31] P. Carvalho, R. P. Paiva, R. Couceiro, J. Henriques, M. Antunes, I. Quintal, J. Muehlsteff, and X. Aubert, “Comparison of systolic time interval measurement modalities for portable devices,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE, Aug. 2010. [Cited on pages iv and 17.]

[32] R. P. Paiva, P. Carvalho, X. Aubert, J. Muehlsteff, J. Henriques, and M. Antunes, “Assessing PEP and LVET from heart sounds: Algorithms and evaluation,” in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, Sep. 2009. [Cited on pages iv and 17.]

[33] S. Janjarasjitt, “Quantification of systolic time intervals using continuous wavelet transform of electrocardiogram and phonocardiogram signals,” in *IFMBE Proceedings*, ser. IFMBE proceedings. Cham: Springer International Publishing, 2020, pp. 356–362. [Cited on pages iv, vi, 17, 19, and 20.]

[34] A. Jimenez-Gonzalez, “Timing the opening and closing of the aortic valve by using the phonocardiogram envelope: a performance test for systolic time intervals measuring,” *Physiol. Meas.*, vol. 42, no. 2, p. 025004, Jan. 2021. [Cited on pages iv, 17, and 21.]

[35] P. Carvalho, R. P. Paiva, R. Couceiro, J. Henriques, I. Quintal, J. Muehlsteff, X. L. Aubert, and M. Antunes, “Assessing systolic time-intervals from heart sound: A feasibility study,” in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, Sep. 2009. [Cited on page 17.]

[36] F. T. Petersen, A. Lobo, C. Oliveira, C. I. Costa, R. Fontes-Carvalho, S. E. Schmidt, and F. Renna, “Predicting the left ventricular ejection fraction using bimodal cardiac auscultation,” 2025, manuscript in preparation. [Cited on pages 25 and 56.]

[37] D. Proaño-Guevara, A. Lobo, C. Oliveira, C. I. Costa, R. Fontes-Carvalho, H. Plácido da Silva, and F. Renna, “Bidirectional fiducial matching of electrocardiography and phonocardiography for multimodal signal quality assessment,” accepted for presentation at the 2025 Computing in Cardiology Conference. [Cited on page 28.]

[38] F. Renna, J. Oliveira, and M. T. Coimbra, “Deep convolutional neural networks for heart sound segmentation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 6, pp. 2435–2445, 2019. [Cited on pages vi, 28, and 30.]

[39] A. Kalyakulina, I. Yusipov, V. Moskalenko, A. Nikolskiy, K. Kosonogov, N. Zolotykh, and M. Ivanchenko, “Lobachevsky University Electrocardiography Database (version 1.0.1),” 2021, rRID:SCR_007345. [Online]. Available: https://doi.org/10.13026/eegm-h675 [Cited on page 31.]

[40] C. Liu, D. Springer, Q. Li, B. Moody, R. Abad, F. Chorro, F. Castells Ramon, J. Roig, I. Silva, A. Johnson, Z. Syed, S. Schmidt, C. Papadaniil, L. Hadjileontiadis, H. Naseri, A. Moukadem, A. Dieterlen, C. Brandt, H. Tang, and G. Clifford, “An open access database for the evaluation of heart sound algorithms,” *Physiological Measurement*, vol. 37, pp. 2181–2213, 11 2016. [Cited on page 31.]

[41] J. Pan and W. J. Tompkins, “A real-time qrs detection algorithm,” *IEEE Transactions on Biomedical Engineering*, vol. 32, no. 3, pp. 230–236, 1985. [Cited on page 37.]

[42] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. H. A. Chen, “NeuroKit2: A python toolbox for neurophysiological signal processing,” *Behavior Research Methods*, vol. 53, no. 4, pp. 1689–1696, 2021. [Cited on page 37.]

[43] P. van Gent, H. Farah, N. van Nes, and B. van Arem, “Heartpy: A novel heart rate algorithm for the analysis of noisy signals,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 66, pp. 368–378, 2019. [Cited on page 37.]

[44] P. Bota, R. Silva, C. Carreiras, A. Fred, and H. P. da Silva, “Biosppy: A python toolbox for physiological signal processing,” *SoftwareX*, vol. 26, p. 101712, 2024. [Cited on page 37.]

[45] G. B. Moody, *WFDB Applications Guide (including GQRS)*, PhysioNet, 2006. [Cited on page 37.]